

M208

Pure mathematics

Book D
Analysis 1

This publication forms part of an Open University module. Details of this and other Open University modules can be obtained from Student Recruitment, The Open University, PO Box 197, Milton Keynes MK7 6BJ, United Kingdom (tel. +44 (0)300 303 5303; email general-enquiries@open.ac.uk).

Alternatively, you may visit the Open University website at www.open.ac.uk where you can learn more about the wide range of modules and packs offered at all levels by The Open University.

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 2018.

Copyright © 2018 The Open University

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, transmitted or utilised in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without written permission from the publisher or a licence from the Copyright Licensing Agency Ltd. Details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Ltd, Barnard's Inn, 86 Fetter Lane, London EC4A 1EN (website www.cla.co.uk).

Open University materials may also be made available in electronic formats for use by students of the University. All rights, including copyright and related rights and database rights, in electronic materials and their contents are owned by or licensed to The Open University, or otherwise used by The Open University as permitted by applicable law.

In using electronic materials and their contents you agree that your use will be solely for the purposes of following an Open University course of study or otherwise as licensed by The Open University or its assigns.

Except as permitted above you undertake not to copy, store in any medium (including electronic storage or use in a website), distribute, transmit or retransmit, broadcast, modify or show in public such electronic materials in whole or in part without the prior written consent of The Open University or in accordance with the Copyright, Designs and Patents Act 1988.

Edited, designed and typeset by The Open University, using L^AT_EX.

Printed in the United Kingdom by Hobbs the Printers Limited, Brunel Road, Totton, Hampshire, SO40 3WX.

ISBN 978 1 4730 2347 5

2.1

Contents

Unit D1 Numbers	1
Introduction to Book D	3
Introduction	3
1 Real numbers	4
1.1 Rational numbers	4
1.2 Decimal representation of rational numbers	5
1.3 Irrational numbers	8
1.4 Real numbers and their properties	9
1.5 Arithmetic with real numbers	12
2 Inequalities	15
2.1 Rearranging inequalities	15
2.2 Solving inequalities	17
2.3 Inequalities involving modulus signs	20
3 Proving inequalities	23
3.1 Triangle Inequality	24
3.2 Proving inequalities by rearrangement	27
3.3 Inequalities involving integers	30
3.4 The Binomial Theorem	31
3.5 Mathematical induction and Bernoulli's Inequality	35
4 Least upper bounds	39
4.1 Upper bounds and lower bounds	39
4.2 Least upper bounds and greatest lower bounds	42
4.3 Least Upper Bound Property	46
5 Manipulating real numbers	49
5.1 Arithmetic with real numbers	49
5.2 Existence of roots	50
5.3 Powers	52
Summary	53
Learning outcomes	54
Solutions to exercises	55

Unit D2 Sequences	61
Introduction	63
1 Introducing sequences	64
1.1 What is a sequence?	64
1.2 Monotonic sequences	67
2 Null sequences	73
2.1 What is a null sequence?	73
2.2 Properties of null sequences	80
2.3 Basic null sequences	86
3 Convergent sequences	88
3.1 What is a convergent sequence?	88
3.2 Combination Rules for convergent sequences	91
3.3 Further rules for convergent sequences	96
4 Divergent sequences	99
4.1 What is a divergent sequence?	100
4.2 Bounded and unbounded sequences	100
4.3 Sequences tending to infinity	102
4.4 Subsequences	106
5 Monotone Convergence Theorem	110
5.1 Convergence of monotonic sequences	110
5.2 The number π	112
5.3 The number e	115
Summary	118
Learning outcomes	118
Solutions to exercises	120

Unit D3 Series	127
Introduction	129
1 Introducing series	129
1.1 What is a convergent series?	129
1.2 Telescoping series	137
1.3 Combination Rules for series	138
1.4 Non-null Test	141
2 Series with non-negative terms	143
2.1 Tests for convergence	144
2.2 Basic series	157
3 Series with positive and negative terms	159
3.1 Absolute convergence	160
3.2 Alternating Test	164
3.3 General strategy	168
4 The exponential function	169
4.1 Definition of e^x	169
4.2 Calculating e	172
4.3 A fundamental property of e^x	174
Summary	176
Learning outcomes	177
Solutions to exercises	178

Unit D4 Continuity	185
Introduction	187
1 Operations on functions	187
1.1 Sums, products and quotients of functions	188
1.2 Composite functions	190
1.3 Inverse functions	192
2 Continuous functions	197
2.1 What is continuity?	198
2.2 Rules for continuous functions	208
2.3 Trigonometric functions and the exponential function	217
3 Properties of continuous functions	221
3.1 Intermediate Value Theorem	222
3.2 Locating zeros of continuous functions	226
3.3 Extreme Value Theorem	230
4 Inverse functions	235
4.1 Inverse Function Rule	235
4.2 Inverses of standard functions	239
4.3 Defining exponential functions	245
4.4 Proof of the Inverse Function Rule	248
Summary	250
Learning outcomes	251
Solutions to exercises	252
Acknowledgements	257
Index	259

Unit D1

Numbers

Introduction to Book D

In this book you will begin your study of the analysis units of this module. *Analysis* is the branch of mathematics that deals in a precise, quantitative way with the concept of a limit, and with the related ideas of infinite sums, continuous functions, differentiation and integration.

The setting for our study of analysis will be real functions – that is, functions whose domains and codomains are subsets of the real line \mathbb{R} . In this book we begin to study such functions from a precise point of view in order to prove many of their properties. You have met many of these properties before, and some may seem intuitively obvious, but this work will provide a sound basis for the study of more difficult properties of functions later in the module.

For example, consider the question:

Does the graph of $y = 2^x$ have a gap at $x = \sqrt{2}$?

The graph in Figure 1 does not appear to have any gaps but we need to check this carefully. Later in the book we answer this question properly by showing that the function $f(x) = 2^x$ has a property called *continuity*, so its graph has no gaps.

Before we can tackle this question, however, there is a preliminary question we must answer:

What precisely is meant by $2^{\sqrt{2}}$?

Thus, to answer a question about real *functions*, we first need to clarify our ideas about real *numbers*.

Introduction

In this unit you will make a deeper study of the *real numbers*, which were introduced in Book A. You will extend your understanding of their properties, and see how they can be represented as infinite decimals. You will meet the rules for manipulating *inequalities*, which play a crucial role in analysis. You will learn how to solve and prove inequalities, and see proofs of several standard results that will be needed in later units. Finally, you will meet the concept of a *least upper bound* and see that the real numbers have the Least Upper Bound Property; this is of great importance in analysis.

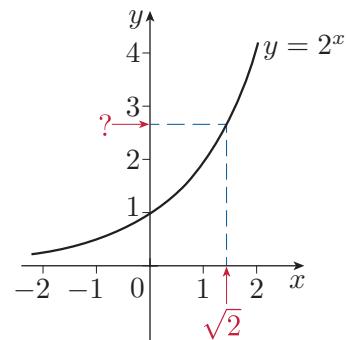


Figure 1 The graph of $y = 2^x$

1 Real numbers

In this section we discuss the real numbers and look at some of their properties. We start by investigating the decimal representations of the rationals, and then proceed to the irrational numbers.

1.1 Rational numbers

The set of **natural numbers** is

$$\mathbb{N} = \{1, 2, 3, \dots\},$$

the set of **integers** is

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$$

and the set of **rational numbers** (or **rationals**) is

$$\mathbb{Q} = \{p/q : p \in \mathbb{Z}, q \in \mathbb{N}\}.$$

(Note that we do not include 0 in the natural numbers, though some mathematical texts do.)

Remember that each rational number has many different representations as a ratio of integers; for example,

$$\frac{1}{3} = \frac{2}{6} = \frac{10}{30} = \dots$$

The usual arithmetic operations of addition, subtraction, multiplication and division can be carried out with rational numbers. We can represent the rationals on a number line as shown in Figure 2.

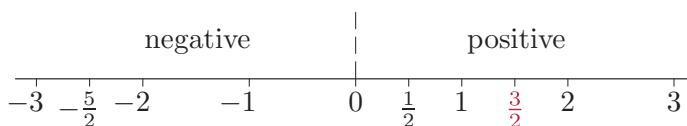


Figure 2 Rational numbers on a number line

For example, the rational $\frac{3}{2}$ is placed at the point which is one half of the way from 0 to 3.

This representation means that rationals have a natural *order* on the number line. For example, $\frac{19}{22}$ lies to the left of $\frac{7}{8}$ because

$$\frac{19}{22} = \frac{76}{88} \quad \text{and} \quad \frac{7}{8} = \frac{77}{88}.$$

If a lies to the left of b on the number line, then we say that

a is *less than* b or b is *greater than* a

and we write

$$a < b \quad \text{or} \quad b > a.$$

For example, we write

$$\frac{19}{22} < \frac{7}{8} \quad \text{or} \quad \frac{7}{8} > \frac{19}{22}.$$

Also, we write $a \leq b$ (or $b \geq a$) if either $a < b$ or $a = b$.

Exercise D1

Arrange the following rationals in order:

$$0, \quad 1, \quad -1, \quad \frac{17}{20}, \quad -\frac{17}{20}, \quad \frac{45}{53}, \quad -\frac{45}{53}.$$

Exercise D2

Show that between any two distinct rationals there is another rational.

1.2 Decimal representation of rational numbers

The decimal system enables us to represent all the natural numbers using only the ten integers

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9,$$

which are called *digits*. We now recall the basic facts about the representation of *rational* numbers by decimals.

Definition

A **decimal** is an expression of the form

$$\pm a_0.a_1a_2a_3\dots,$$

where a_0 is a non-negative integer and a_n is a digit for each $n \in \mathbb{N}$.

For example, $13.1212\dots$ and $-1.111\dots$ are both decimals. If only a finite number of the digits a_1, a_2, \dots are non-zero, then the decimal is a **terminating** or **finite decimal**, and we usually omit the tail of zeros. For example, we usually write 0.85 instead of $0.8500\dots$

Terminating decimals are used to represent rational numbers in the following way:

$$\pm a_0.a_1a_2\dots a_n = \pm \left(a_0 + \frac{a_1}{10^1} + \frac{a_2}{10^2} + \dots + \frac{a_n}{10^n} \right).$$

For example,

$$0.85 = 0 + \frac{8}{10^1} + \frac{5}{10^2} = \frac{85}{100} = \frac{17}{20}.$$

It can be shown that any fraction for which the factors of the denominator are all powers of 2 and/or 5 (for example, any fraction with denominator $20 = 2^2 \times 5$) can be represented by such a terminating decimal, which can be found by long division; and conversely, that every terminating decimal represents such a fraction.

However, if we apply long division to other rationals, then the process of long division never terminates and we obtain a **non-terminating** or **infinite decimal**. For example, applying long division to $1/3$ gives $0.333\dots$ and for $19/22$ we obtain $0.863\,63\dots$.

Exercise D3

Use long division to find the decimal corresponding to $1/7$.

The infinite decimals obtained by applying the long division process to rationals have a certain common property. All of them are **recurring decimals**; that is, they have a repeating block of digits, so they can be written in shorthand form by placing a line over the repeating block, as follows:

$$\begin{array}{ll} 0.333\dots & = 0.\overline{3}, \\ 0.142\,857\,142\,857\dots & = 0.\overline{142\,857}, \\ 0.863\,63\dots & = 0.\overline{863}. \end{array}$$

(Another commonly used notation for recurring decimals is to place a dot over the first and last digits in the repeating block; for example,

$$0.\dot{3} \quad \text{and} \quad 0.1\dot{4}2\dot{8}\dot{5}\dot{7}.$$

However, we do not use this notation in this module.)

To see why applying the long division process to a fraction p/q always leads to either a recurring decimal or a terminating decimal, note that there are only q possible remainders at each stage of the division, so one of these remainders must eventually repeat. When this happens, the block of digits obtained after the first occurrence of this remainder will be repeated infinitely often. If the remainder 0 occurs, then the resulting decimal is a terminating decimal; that is, it ends in recurring 0s.

Recurring decimals which arise from the long division of fractions are used to represent the corresponding rational numbers. Conversely, it can be shown that every recurring decimal represents some rational number.

However, the representation of rational numbers by recurring decimals is not quite as straightforward as for terminating decimals. If we try the same approach, we get an equation involving the sum of infinitely many terms, for example,

$$\frac{1}{3} = 0.\overline{3} = \frac{3}{10^1} + \frac{3}{10^2} + \frac{3}{10^3} + \dots,$$

and it is not immediately clear what such a sum means. This will be made

precise when you have met the idea of the sum of a convergent infinite series later in the book. For the moment, though, when we write the statement $1/3 = 0.\overline{3}$, we mean simply that the decimal $0.\overline{3}$ arises from $1/3$ by the long division process.

The following worked exercise illustrates one way of finding the fraction with a given decimal representation.

Worked Exercise D1

Find the fraction whose decimal representation is $0.8\overline{63}$.

Solution

 We begin by finding the fraction whose decimal representation is equal to $0.\overline{63}$. 

Let $x = 0.\overline{63}$.

 Because the recurring block has length two, we multiply both sides by 10^2 . 

Multiplying both sides by 10^2 , we obtain

$$100x = 63.\overline{63} = 63 + x.$$

Hence

$$99x = 63, \quad \text{so} \quad x = \frac{63}{99} = \frac{7}{11}.$$

 Having found x as a fraction, we now write $0.8\overline{63}$ as the sum of two fractions. 

Thus

$$0.8\overline{63} = \frac{8}{10} + \frac{x}{10} = \frac{8}{10} + \frac{7}{110} = \frac{95}{110} = \frac{19}{22}.$$

The key idea in the solution above is that multiplication of a decimal by 10^k , where $k \in \mathbb{N}$, moves the decimal point k places to the right.

Exercise D4

Using the method of Worked Exercise D1, find the fractions represented by the following decimals.

- (a) $0.\overline{231}$ (b) $2.2\overline{81}$

Decimals which end in recurring 9s sometimes arise as alternative representations for terminating decimals. For example,

$$1 = 0.\overline{9} = 0.999\dots \quad \text{and} \quad 1.35 = 1.34\overline{9} = 1.3499\dots$$

You may find this rather disconcerting, but it is important to realise that this representation is a matter of *definition*. We wish to allow the decimal $0.999\dots$ to represent a number x . This number x must be less than or equal to 1 and greater than each of the numbers

$$0.9, 0.99, 0.999, \dots$$

The *only* rational with these properties is 1. When possible, we avoid using the form of a decimal which ends in recurring 9s.

The decimal representation of rational numbers has the advantage that it enables us to decide immediately which of two distinct positive rationals is the greater. We need only examine their decimal representations and notice the first place at which the digits differ. For example, to order $7/8$ and $19/22$, we can write

$$\frac{7}{8} = 0.875 \quad \text{and} \quad \frac{19}{22} = 0.863\overline{63}\dots$$

Then

$$\begin{array}{c} \downarrow \\ 0.863\overline{63}\dots < 0.875, \end{array} \quad \text{so} \quad \frac{19}{22} < \frac{7}{8}.$$

Note that, when comparing decimals in this way, the decimal on the left should not end in recurring 9s unless the decimal on the right is also non-terminating.

Exercise D5

Find the first two digits after the decimal point in the decimal representations of $17/20$ and $45/53$, and hence determine which of these two rationals is the greater.

1.3 Irrational numbers

You saw in Unit A2 *Number systems* that there is no rational number which satisfies the equation $x^2 = 2$; that is, $\sqrt{2}$ is not rational.

There are many other mathematical quantities that cannot be described exactly by rational numbers. For example:

- if m and n are natural numbers, and the equation $x^m = n$ has no integer solution, then the positive solution of this equation, written as $\sqrt[m]{n}$, is not rational
- the number π , which denotes the ratio of the circumference of a circle to its diameter
- the number e , the base of natural logarithms.

A number which is not rational is called **irrational**. It is natural to ask whether irrational numbers, such as $\sqrt{2}$ and π , can be represented as decimals. Using your calculator, you can check that

$$(1.414\ 213\ 56)^2$$

is very close to 2, so 1.414 213 56 is a very good approximate value for $\sqrt{2}$. But is there a decimal that represents $\sqrt{2}$ exactly?

In fact, it is possible to represent all irrational numbers by **non-recurring decimals**, that is, infinite decimals that do not end in a recurring block of digits. For example, there are non-recurring decimals representing $\sqrt{2}$ and π , the first few digits of which are

$$\sqrt{2} = 1.414\ 213\ 56\dots \quad \text{and} \quad \pi = 3.141\ 592\ 65\dots$$

Conversely, it is also natural to ask whether arbitrary non-recurring decimals, such as

$$0.101\ 001\ 000\ 100\ 001\dots \quad \text{and} \quad 0.123\ 456\ 789\ 101\ 112\dots,$$

always represent irrational numbers. We take it as a basic assumption about the number system that they do.

Thus the set of irrational numbers consists of all the non-recurring decimals.

1.4 Real numbers and their properties

We can now define what we mean by the set of real numbers.

Definition

The set of **real numbers**, denoted by \mathbb{R} , is the union of the set of rational numbers and the set of irrational numbers. In other words, it is the set of all terminating, recurring and non-recurring decimals.

As with rational numbers, we can determine which of two real numbers is greater by comparing their decimals and noticing the first pair of corresponding digits which differ. For example,

$$\begin{array}{c} \downarrow & \downarrow \\ 0.101\ 001\ 000\ 100\ 001\dots & < 0.123\ 456\ 789\ 101\ 112\dots \end{array}$$

We now use this idea to associate with each irrational number a point on the number line. For example, as illustrated in Figure 3, the irrational number whose decimal representation begins

$$x = 0.123456789101112\dots$$

satisfies each of the inequalities

$$\begin{aligned} 0.1 &< x < 0.2 \\ 0.12 &< x < 0.13 \\ 0.123 &< x < 0.124 \\ &\vdots \end{aligned}$$



Figure 3 The point on the number line corresponding to $x = 0.12345\dots$

We assume that there is a point on the number line corresponding to x which lies to the right of each of the rational numbers $0.1, 0.12, 0.123, \dots$ and to the left of each of the rational numbers $0.2, 0.13, 0.124, \dots$.

As usual, negative real numbers correspond to points lying to the left of 0. The number line, complete with both rational and irrational points, is called the **real line**; see Figure 4.

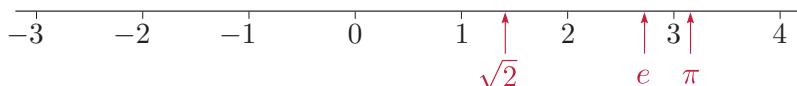


Figure 4 The real line

Our definition of the real numbers is therefore consistent with the picture of the real numbers as a number line, which you met in Unit A2. Saying the same thing more formally, there is a one-to-one correspondence between the points on the real line and the set \mathbb{R} of real numbers, as defined above.

In the box below, we list several *order properties* of \mathbb{R} . You are probably already familiar with these, though you may not have met their names before.

Order properties of \mathbb{R}

Trichotomy Property If $a, b \in \mathbb{R}$, then *exactly one* of the following holds:

$$a < b \quad \text{or} \quad a = b \quad \text{or} \quad a > b.$$

Transitive Property If $a, b, c \in \mathbb{R}$, then

$$a < b \text{ and } b < c \implies a < c.$$

Archimedean Property If $a \in \mathbb{R}$, then there is a positive integer n such that

$$n > a.$$

Density Property If $a, b \in \mathbb{R}$ and $a < b$, then there is a rational number x and an irrational number y such that

$$a < x < b \quad \text{and} \quad a < y < b.$$

These properties are used frequently in analysis, though they are not often referred to explicitly by name. In this unit, however, we often use the names to point out when the properties are being used, to aid your understanding.

Each of the order properties in the above box can be proved from our definition of the real numbers, but we do not give the proofs here. The first three are almost self-evident, but the Density Property is not so obvious. One consequence of the Density Property is that between any two distinct real numbers there are infinitely many rational numbers and infinitely many irrational numbers. The next worked exercise gives an example of how to find a rational number and an irrational number between two given real numbers. The method is indicative of how the Density Property could be proved for general real numbers a and b .

Worked Exercise D2

Let $a = 0.1\bar{2}$ and $b = 0.12345\dots$. Find a rational number x and an irrational number y such that

$$a < x < b \quad \text{and} \quad a < y < b.$$

Solution

There are many different methods you could use. One method is to begin by noting that the two decimals

$$\begin{array}{ccc} \downarrow & & \downarrow \\ a = 0.12333\dots & \text{and} & b = 0.12345\dots \end{array}$$

differ first at the fourth digit after the decimal point. If we truncate b after this digit, then we obtain a suitable rational number.

For example, the rational number

$$x = 0.1234$$

satisfies $a < x < b$.

To find an irrational number y between a and b , we attach to x a (sufficiently small) non-recurring tail, such as $010010001\dots$. The resulting number is irrational since its decimal representation is non-recurring.

The irrational number

$$y = 0.1234 \left| \begin{array}{l} 010010001\dots \\ \text{non-recurring tail} \end{array} \right.$$

satisfies $a < y < b$.

Exercise D6

Let $a = 0.\bar{3}$ and $b = 0.3401$. Find a rational number x and an irrational number y such that $a < x < b$ and $a < y < b$.

1.5 Arithmetic with real numbers

We now turn to the question of how to carry out arithmetic using real numbers, bearing in mind our definition of the real numbers in the last subsection.

You saw in Unit A2 that the set \mathbb{R} of real numbers, in common with the set \mathbb{Q} of rational numbers, forms a *field*. This means that the properties in the box below hold for arithmetic in \mathbb{R} .

Arithmetic in \mathbb{R}

Properties for addition

A1 Closure For all $a, b \in \mathbb{R}$,

$$a + b \in \mathbb{R}.$$

A2 Associativity For all $a, b, c \in \mathbb{R}$,

$$a + (b + c) = (a + b) + c.$$

A3 Additive identity For all $a \in \mathbb{R}$,

$$a + 0 = a = 0 + a.$$

A4 Additive inverses For each $a \in \mathbb{R}$, there is a number $-a \in \mathbb{R}$ such that

$$a + (-a) = 0 = (-a) + a.$$

A5 Commutativity For all $a, b \in \mathbb{R}$,

$$a + b = b + a.$$

Properties for multiplication

M1 Closure For all $a, b \in \mathbb{R}$,

$$a \times b \in \mathbb{R}.$$

M2 Associativity For all $a, b, c \in \mathbb{R}$,

$$a \times (b \times c) = (a \times b) \times c.$$

M3 Multiplicative identity For all $a \in \mathbb{R}$,

$$a \times 1 = a = 1 \times a.$$

M4 Multiplicative inverses For each $a \in \mathbb{R}^*$, there is a number $a^{-1} \in \mathbb{R}$ such that

$$a \times a^{-1} = 1 = a^{-1} \times a.$$

M5 Commutativity For all $a, b \in \mathbb{R}$,

$$a \times b = b \times a.$$

Property combining addition and multiplication

D1 Distributivity For all $a, b, c \in \mathbb{R}$,

$$a \times (b + c) = a \times b + a \times c.$$

Put more succinctly, the properties in the box mean that:

- \mathbb{R} is an abelian group under the operation of addition $+$
- $\mathbb{R}^* = \mathbb{R} - \{0\}$ is an abelian group under the operation of multiplication \times ;

and these two group structures are linked by the distributive property.

In Unit A2, these properties were introduced simply as a way of formalising the elementary rules of arithmetic that were already familiar to you. However, now that we have defined the real numbers as the set of all terminating, recurring and non-recurring decimals, we need to show that these properties follow from the definition.

For terminating and recurring decimals (that is, the rational numbers), this is straightforward: we can do arithmetic by first converting the decimals to the corresponding fractions. However, it is not obvious how to do arithmetic with non-recurring decimals (that is, irrationals). For example, assuming that we can represent $\sqrt{2}$ and π by the non-recurring decimals that begin

$$\sqrt{2} = 1.41421356\dots \quad \text{and} \quad \pi = 3.14159265\dots,$$

can we also represent the sum $\sqrt{2} + \pi$ and the product $\sqrt{2} \times \pi$ as decimals? In other words, what is meant by the operations of addition and multiplication when non-recurring decimals are involved, and do these operations satisfy the properties stated in the above box?

It turns out that setting up the appropriate definitions and proving the necessary properties is a lengthy process, and we will not go into this here. The important point is that, using our definition of the real numbers, addition and multiplication *can* be formally defined, and it can be proved that all the above properties hold in \mathbb{R} . From now on, we *assume* that this process has been carried out. Furthermore, we assume that the set \mathbb{R} contains the n th roots and rational powers of positive real numbers, with their usual properties. We describe one way to justify these assumptions in Section 5.

We conclude this section by noting that analysis texts take various approaches to defining the real numbers. For example, it is common to assume that there exists a set \mathbb{R} which is a field containing \mathbb{Q} and having certain extra properties, and then to deduce all results from these assumptions. In this ‘axiomatic approach’ the definition of the real numbers themselves may not be given (though they can be defined by a somewhat abstract procedure involving partitions of \mathbb{Q} called ‘Dedekind cuts’), but it is then *proved* that each real number must have a decimal representation. In this module, we adopt a more concrete approach in which the real numbers are *defined* to be decimals.

2 Inequalities

Much of analysis is concerned with inequalities of various kinds; the aim of this section and the next is to provide practice in their manipulation. In this section you will meet the rules for manipulating inequalities and see how to solve inequalities by using these rules.

2.1 Rearranging inequalities

The fundamental rule, on which much manipulation of inequalities is based, is that the statement $a < b$ means exactly the same as the statement $b - a > 0$. We express this as follows.

Rule 1

$$a < b \iff b - a > 0.$$

Recall that the symbol \iff means ‘if and only if’ or ‘is equivalent to’. Thus, put another way, this rule says that the inequalities $a < b$ and $b - a > 0$ are equivalent.

There are several other standard rules for rearranging an inequality into an equivalent form. Each of these can be deduced from Rule 1, although the proofs are not given here. For example, we obtain an equivalent inequality by adding the same expression to both sides.

Rule 2

$$a < b \iff a + c < b + c.$$

Another way to rearrange an inequality is to multiply both sides by an expression that is strictly greater than zero. It is also possible to multiply both sides by an expression that is strictly less than zero, but in this case the inequality must be reversed. For example, $2 < 3 \iff -2 > -3$.

Rule 3

If $c > 0$, then

$$a < b \iff ac < bc;$$

if $c < 0$, then

$$a < b \iff ac > bc.$$

Sometimes the most effective way to rearrange an inequality is to take reciprocals. You can do this if both sides of the inequality are positive, and it is important to remember that the direction of the inequality has to be reversed. For example, $2 < 3 \iff \frac{1}{2} > \frac{1}{3}$.

Rule 4

If $a, b > 0$, then

$$a < b \iff \frac{1}{a} > \frac{1}{b}.$$

Some inequalities can be simplified only by taking powers. In order to do this, both sides must be non-negative and the power must be positive.

Rule 5

If $a, b \geq 0$ and $p > 0$, then

$$a < b \iff a^p < b^p.$$

For positive integers p , Rule 5 follows from the identity

$$b^p - a^p = (b - a)(b^{p-1} + b^{p-2}a + \cdots + ba^{p-2} + a^{p-1});$$

since the value of the right-hand bracket is positive, we deduce that

$$b - a > 0 \iff b^p - a^p > 0.$$

For other positive real numbers p the proof of Rule 5 is harder, but the rule remains true.

There are corresponding versions of Rules 1–5 in which the *strict* inequality $a < b$ is replaced by the *weak* inequality $a \leq b$. For example, if $c > 0$, then

$$a \leq b \iff ac \leq bc.$$

The box below summarises the rules you have met so far for manipulating inequalities.

Rules for rearranging inequalities

Let a, b, c and p be real numbers.

Rule 1 $a < b \iff b - a > 0$.

Rule 2 $a < b \iff a + c < b + c$.

Rule 3 If $c > 0$, then $a < b \iff ac < bc$;
if $c < 0$, then $a < b \iff ac > bc$.

Rule 4 If $a, b > 0$, then

$$a < b \iff \frac{1}{a} > \frac{1}{b}.$$

Rule 5 If $a, b \geq 0$ and $p > 0$, then

$$a < b \iff a^p < b^p.$$

All the above rules also hold if strict inequalities are replaced by weak inequalities.

In manipulating inequalities, we also make frequent use of the usual rules for the sign of a product:

\times	+	-
+	+	-
-	-	+

In particular, the square of any real number is non-negative.

The next exercise gives you a chance to practise using the rules for rearranging inequalities.

Exercise D7

In each of the following cases, apply the rules to the inequality $x > 2$ to obtain an equivalent inequality that contains the given expression, noting carefully which rules you are using.

- (a) $x + 3$ (b) $2 - x$ (c) $5x + 2$ (d) $1/(5x + 2)$

2.2 Solving inequalities

Solving an inequality that involves an unknown real number x means determining the values of x for which the inequality is true; that is, finding the **solution set** of the inequality, usually as a union of intervals. We can often do this by rewriting the inequality in an equivalent but simpler form, using Rules 1–5.

Worked Exercise D3

Solve the inequality

$$\frac{x+2}{x+4} > \frac{x-3}{2x-1}.$$

Solution

Observe that we *cannot* solve this inequality by cross-multiplying (that is, by multiplying both sides by the product of both denominators), because the denominators have different signs or are zero for some values of x , so we are unable to apply Rule 3. Instead, we rearrange the inequality using Rule 1, to give an equivalent inequality with just 0 on one side.

Rearranging the inequality gives

$$\begin{aligned} \frac{x+2}{x+4} > \frac{x-3}{2x-1} &\iff \frac{x+2}{x+4} - \frac{x-3}{2x-1} > 0 \\ &\iff \frac{x^2 + 2x + 10}{(x+4)(2x-1)} > 0. \end{aligned}$$

By completing the square, we obtain

$$x^2 + 2x + 10 = (x + 1)^2 + 9,$$

so the numerator is always positive.

We can now find the solution set using a table of signs:

x	$(-\infty, -4)$	-4	$(-4, \frac{1}{2})$	$\frac{1}{2}$	$(\frac{1}{2}, \infty)$
$x^2 + 2x + 10$	+	+	+	+	+
$x + 4$	-	0	+	+	+
$2x - 1$	-	-	-	0	+
$\frac{x^2 + 2x + 10}{(x + 4)(2x - 1)}$	+	*	-	*	+

So the solution set is

$$\left\{ x : \frac{x+2}{x+4} > \frac{x-3}{2x-1} \right\} = (-\infty, -4) \cup (\frac{1}{2}, \infty).$$

We now consider an inequality that *could* be solved by using cross-multiplication, though in this case it is easier to take reciprocals using Rule 4.

Worked Exercise D4

Solve the inequality

$$\frac{1}{2x^2 + 2} < \frac{1}{4}.$$

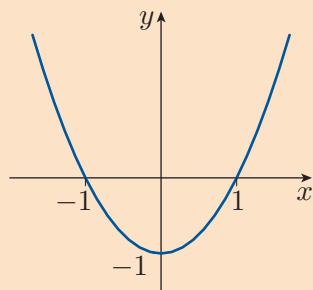
Solution

Since both sides of the inequality are always positive, we can apply Rule 4.

Since $2x^2 + 2 > 0$, we can rearrange the inequality into an equivalent form as follows:

$$\begin{aligned} \frac{1}{2x^2 + 2} < \frac{1}{4} &\iff 2x^2 + 2 > 4 \quad (\text{by Rule 4}) \\ &\iff x^2 + 1 > 2 \quad (\text{by Rule 3}) \\ &\iff x^2 - 1 > 0 \quad (\text{by Rule 1}) \\ &\iff (x - 1)(x + 1) > 0. \end{aligned}$$

At this point, it might be helpful to write down a table of signs or, alternatively, quickly sketch the parabola that is the graph of $y = (x - 1)(x + 1)$.



So the solution set is

$$\left\{ x : \frac{1}{2x^2 + 2} < \frac{1}{4} \right\} = (-\infty, -1) \cup (1, \infty).$$

Exercise D8

Solve the following inequalities.

(a) $\frac{4x - x^2 - 7}{x^2 - 1} \geq 3$ (b) $2x^2 \geq (x + 1)^2$

We now consider an inequality which involves rational powers. Here we need to be careful, when applying Rule 5, to ensure that both sides of the inequality are non-negative.

Worked Exercise D5

Solve the inequality

$$\sqrt{2x + 3} > x.$$

Solution

We can get rid of the awkward square root sign in the inequality by squaring both sides – that is, by applying Rule 5 with $p = 2$. However, we can do this only if both sides are non-negative.

Remember that $\sqrt{}$ always means the non-negative square root.

The expression $\sqrt{2x + 3}$ is defined only when $2x + 3 \geq 0$, that is, for $x \geq -3/2$. Hence we need only consider those x in $[-3/2, \infty)$.

For $x \geq 0$, we have

$$\begin{aligned} \sqrt{2x + 3} > x &\iff 2x + 3 > x^2 \quad (\text{by Rule 5, with } p = 2) \\ &\iff x^2 - 2x - 3 < 0 \quad (\text{by Rule 1}) \\ &\iff (x - 3)(x + 1) < 0. \end{aligned}$$

So the part of the solution set in $[0, \infty)$ is $[0, 3)$.

On the other hand, if $-3/2 \leq x < 0$, then

$$\sqrt{2x+3} \geq 0 > x,$$

so the other part of the solution set is $[-3/2, 0)$.

Hence the complete solution set is

$$\{x : \sqrt{2x+3} > x\} = [-3/2, 0) \cup [0, 3) = [-3/2, 3).$$

Exercise D9

Solve the inequality

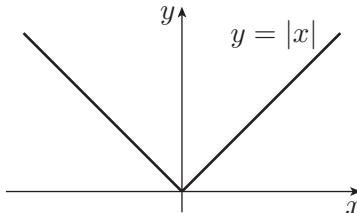
$$\sqrt{2x^2 - 2} > x.$$

2.3 Inequalities involving modulus signs

Now we consider inequalities involving the *modulus* of a real number. Recall that if $a \in \mathbb{R}$, then its **modulus**, or **absolute value**, $|a|$ is defined by

$$|a| = \begin{cases} a, & \text{if } a \geq 0, \\ -a, & \text{if } a < 0. \end{cases}$$

Figure 5 The graph of $y = |x|$



The graph of $y = |x|$ is illustrated in Figure 5.

It is useful to think of $|a|$ as the distance along the real line from 0 to a . For example, $|3| = |-3| = 3$ as shown in Figure 6.

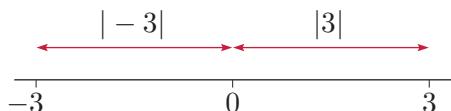


Figure 6 The modulus of 3 and -3

In the same way, $|a - b|$ is the distance from 0 to $a - b$, which is the same as the distance from a to b , as illustrated in Figure 7. For example, the distance from -2 to 3 is $|(-2) - 3| = |-5| = 5$.

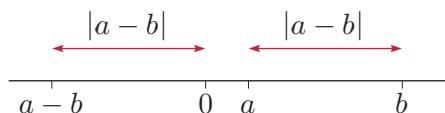


Figure 7 The modulus of $a - b$

Note also that $|a + b| = |a - (-b)|$ is the distance from a to $-b$, as illustrated in Figure 8.

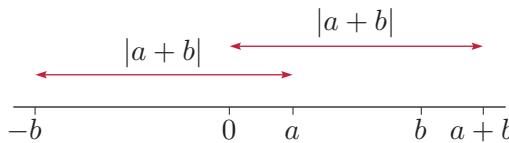


Figure 8 The modulus of $a + b$

We list below some basic properties of the modulus, which follow immediately from the definition.

Properties of the modulus

If $a, b \in \mathbb{R}$, then

1. $|a| \geq 0$, with equality if and only if $a = 0$
2. $-|a| \leq a \leq |a|$
3. $|a|^2 = a^2$
4. $|a - b| = |b - a|$
5. $|ab| = |a||b|$.

We now give a rule for rearranging inequalities that involve the modulus of an expression. The rule follows from property 2 in the above list, and as usual there is a corresponding version with weak inequalities.

Rule 6

$$|a| < b \iff -b < a < b.$$

You can use this rule alongside the other rules for rearranging inequalities that you met in Subsection 2.1. For reference, here is a summary of all the rules.

Rules for rearranging inequalities

Let a, b, c and p be real numbers.

Rule 1 $a < b \iff b - a > 0$.

Rule 2 $a < b \iff a + c < b + c$.

Rule 3 If $c > 0$, then $a < b \iff ac < bc$;
if $c < 0$, then $a < b \iff ac > bc$.

Rule 4 If $a, b > 0$, then

$$a < b \iff \frac{1}{a} > \frac{1}{b}.$$

Rule 5 If $a, b \geq 0$ and $p > 0$, then

$$a < b \iff a^p < b^p.$$

Rule 6 $|a| < b \iff -b < a < b$.

All the above rules also hold if strict inequalities are replaced by weak inequalities.

Worked Exercise D6

Solve the inequality $|x - 2| < 1$.

Solution

We rearrange the inequality into an equivalent form by using Rule 6 with $a = x - 2$ and $b = 1$.

We have

$$\begin{aligned} |x - 2| < 1 &\iff -1 < x - 2 < 1 \quad (\text{by Rule 6}) \\ &\iff 1 < x < 3. \end{aligned}$$

So the solution set is

$$\{x : |x - 2| < 1\} = (1, 3).$$

We can also rearrange inequalities involving modulus signs by using Rule 5 with $p = 2$, since the modulus of a number is non-negative.

Worked Exercise D7

Solve the inequality $|x - 2| \leq |x + 1|$.

Solution

We rearrange the inequality into an equivalent form by using Rule 5 with $a = |x - 2|$, $b = |x + 1|$ and $p = 2$.

We have

$$\begin{aligned} |x - 2| \leq |x + 1| &\iff (x - 2)^2 \leq (x + 1)^2 \quad (\text{by Rule 5, with } p = 2) \\ &\iff x^2 - 4x + 4 \leq x^2 + 2x + 1 \\ &\iff 3 \leq 6x \\ &\iff \frac{1}{2} \leq x. \end{aligned}$$

So the solution set is

$$\{x : |x - 2| \leq |x + 1|\} = [\frac{1}{2}, \infty).$$

Thinking about what an inequality means geometrically can often give you an idea of its solution set.

For example, the inequalities in Worked Exercises D6 and D7 can be interpreted geometrically as follows.

The inequality

$$|x - 2| < 1$$

holds when the distance from x to 2 is strictly less than 1. So it holds when x lies in the open interval $(1, 3)$, which has midpoint 2. This is what we found in Worked Exercise D6 by using Rule 6.

Similarly, the inequality

$$|x - 2| \leq |x + 1|$$

holds when the distance from x to 2 is less than or equal to the distance from x to -1 , since $|x + 1| = |x - (-1)|$. As the point halfway from -1 to 2 is $\frac{1}{2}$, the inequality holds when x lies in the interval $[\frac{1}{2}, \infty)$, as illustrated in Figure 9. This agrees with what we found algebraically in Worked Exercise D7.

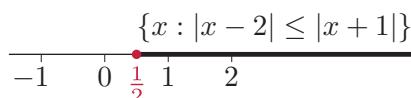


Figure 9 The solution set of $|x - 2| \leq |x + 1|$

Exercise D10

Solve the following inequalities.

- (a) $|2x^2 - 13| < 5$ (b) $|x - 1| \leq 2|x + 1|$

3 Proving inequalities

In this section you will see how to *prove* inequalities of various types.

Several of the inequalities we prove here will be used in later analysis units of this module.

In proving inequalities, we will make use of the rules for rearranging inequalities that you met in Section 2, together with some further rules for deducing ‘new inequalities from old’, which we outline here.

You met the first such rule in Section 1, where it was called the Transitive Property of \mathbb{R} .

Transitive Rule for inequalities

$$a < b \text{ and } b < c \implies a < c.$$

We use the Transitive Rule when we want to prove that $a < c$ and we know that $a < b$ and $b < c$.

The following rules are also useful.

Combination Rules for inequalities

If $a < b$ and $c < d$, then

Sum Rule $a + c < b + d$

Product Rule $ac < bd$, provided $a, c \geq 0$.

For example, since $2 < 3$ and $4 < 5$, it follows that

$$2 + 4 < 3 + 5,$$

$$2 \times 4 < 3 \times 5.$$

There are versions of the Transitive Rule and Combination Rules involving weak inequalities. For example, a weak version of the Transitive Rule is

$$a \leq b \text{ and } b \leq c \implies a \leq c.$$

Notice that the Transitive Rule and the Combination Rules have a different nature from Rules 1–6 given in Section 2. Rules 1–6 tell us how to rearrange inequalities into *equivalent* inequalities, whereas the Transitive Rule and the Combination Rules enable us to start with two inequalities and deduce a new inequality which is *not* equivalent to either of the old ones, but does follow from them. For example, if we know that $x < y$ and $y < 5$, then we can use the Transitive Rule to deduce that $x < 5$; this new inequality is not equivalent to either of the inequalities that we started with, but it does follow from them.

3.1 Triangle Inequality

Now we meet an inequality that can be used to deduce ‘new inequalities from old’, but is also of great importance in its own right.

This inequality involves the modulus of three real numbers a , b and $a + b$, and is called the *Triangle Inequality* because it is related to the fact that the length of one side of a triangle is less than the sum of the lengths of the other two sides. As you will see, the Triangle Inequality has many applications in analysis.

Triangle Inequality

If $a, b \in \mathbb{R}$, then

1. $|a + b| \leq |a| + |b|$ (usual form)
2. $|a - b| \geq ||a| - |b||$ ('backwards' form).

Proof

1. We rearrange the inequality into an equivalent form:

$$\begin{aligned} |a+b| \leq |a| + |b| &\iff (a+b)^2 \leq (|a|+|b|)^2 \quad (\text{by Rule 5, with } p=2) \\ &\iff a^2 + 2ab + b^2 \leq a^2 + 2|a||b| + b^2 \\ &\iff 2ab \leq |2ab|. \end{aligned}$$

This final inequality is certainly true for all $a, b \in \mathbb{R}$, so the first inequality must also be true for all $a, b \in \mathbb{R}$.

2. We prove the backwards form using the same method:

$$\begin{aligned} |a-b| \geq ||a|-|b|| &\iff (a-b)^2 \geq (|a|-|b|)^2 \quad (\text{by Rule 5, with } p=2) \\ &\iff a^2 - 2ab + b^2 \geq a^2 - 2|a||b| + b^2 \\ &\iff -2ab \geq -|2ab| \\ &\iff 2ab \leq |2ab|, \end{aligned}$$

which, as before, is true for all $a, b \in \mathbb{R}$. ■

Remarks

1. Although we have used implications in both directions here, the proof requires only the implications going from right to left. For example, in the proof of the usual form of the Triangle Inequality, the important implication is

$$|a+b| \leq |a| + |b| \iff 2ab \leq |2ab|.$$

2. The usual form of the Triangle Inequality can also be proved by using Rule 6, which gives

$$|a+b| \leq |a| + |b| \iff -(|a| + |b|) \leq a+b \leq |a| + |b|. \quad (1)$$

Now since we know that

$$-|a| \leq a \leq |a| \quad \text{and} \quad -|b| \leq b \leq |b|,$$

it follows from the Sum Rule that

$$-(|a| + |b|) \leq a+b \leq |a| + |b|,$$

which is the statement on the right in (1). Hence the equivalent statement on the left in (1) also holds, and this is the Triangle Inequality:

$$|a+b| \leq |a| + |b|.$$

3. There is a version of the Triangle Inequality for n real numbers, where $n \geq 2$:

$$|a_1 + a_2 + \cdots + a_n| \leq |a_1| + |a_2| + \cdots + |a_n|.$$

This can be proved using the Principle of Mathematical Induction, which you met in Book A and will use again in Subsection 3.5 of this unit.

The worked exercise below illustrates some typical applications of the Triangle Inequality. Remember that, in each part, we are deducing one inequality from another, *not* showing that two inequalities are equivalent.

Worked Exercise D8

Use the Triangle Inequality to prove the following statements.

$$(a) |a| \leq 1 \implies |3 + a^3| \leq 4 \quad (b) |b| < 1 \implies |3 - b| > 2$$

Solution

(a) Suppose that $|a| \leq 1$. The Triangle Inequality gives

$$\begin{aligned} |3 + a^3| &\leq |3| + |a^3| \\ &= 3 + |a|^3. \end{aligned}$$

Now $|a| \leq 1$ and therefore

$$3 + |a|^3 \leq 3 + 1 = 4.$$

We deduce, using the Transitive Rule, that

$$|a| \leq 1 \implies |3 + a^3| \leq 4.$$

(b) Suppose that $|b| < 1$. The backwards form of the Triangle Inequality gives

$$\begin{aligned} |3 - b| &\geq ||3| - |b|| \\ &= |3 - |b|| \\ &\geq 3 - |b|. \end{aligned}$$

Now $|b| < 1$, so $-|b| > -1$, and hence

$$3 - |b| > 3 - 1 = 2.$$

We deduce, using the previous chain of inequalities, that

$$|b| < 1 \implies |3 - b| > 2.$$

Remarks

1. The statements proved in Worked Exercise D8 can also be written in the following form:

$$|3 + a^3| \leq 4, \quad \text{for } |a| \leq 1$$

and

$$|3 - b| > 2, \quad \text{for } |b| < 1.$$

2. Notice that the reverse implications

$$|3 + a^3| \leq 4 \implies |a| \leq 1 \quad \text{and} \quad |3 - b| > 2 \implies |b| < 1$$

are *false*. For example, try putting $a = -\frac{3}{2}$ and $b = -2$.

Exercise D11

Use the Triangle Inequality to prove the following statements.

$$(a) |a| \leq \frac{1}{2} \implies |a+1| \leq \frac{3}{2} \quad (b) |b| < \frac{1}{2} \implies |b^3 - 1| > \frac{7}{8}$$

3.2 Proving inequalities by rearrangement

We now prove some further inequalities using the method you saw in the proof of the Triangle Inequality. As in that proof, we start from the inequality we wish to prove, and use the various rearrangement rules to obtain a chain of equivalent inequalities until we reach an inequality that we know must be true.

In the worked exercises in this subsection, we apply this technique to prove two inequalities with geometric interpretations. The first states that

$$ab \leq \left(\frac{a+b}{2}\right)^2, \quad \text{for } a, b \in \mathbb{R}.$$

This inequality has the geometric interpretation illustrated in Figure 10: the area of a rectangle with sides of length a and b is less than or equal to the area of a square with the same perimeter, that is, with sides of length $(a+b)/2$.

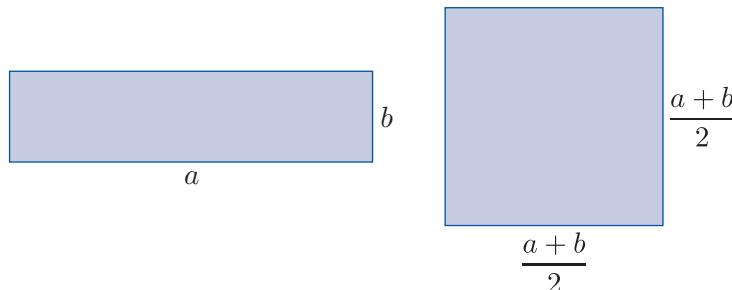


Figure 10 Comparison between the area of a rectangle and the area of a square with the same perimeter

In the worked exercises we include some comments about which of the rearrangement rules we are using, but you need not do this when you write out solutions.

Worked Exercise D9

Prove that

$$ab \leq \left(\frac{a+b}{2}\right)^2, \quad \text{for } a, b \in \mathbb{R}.$$

Solution

We begin by multiplying out the bracket to see if this helps to simplify the inequality. We then apply the rearrangement rules.

Rearranging the inequality, we obtain:

$$\begin{aligned} ab \leq \left(\frac{a+b}{2}\right)^2 &\iff ab \leq \frac{a^2 + 2ab + b^2}{4} \\ &\iff 4ab \leq a^2 + 2ab + b^2 && (\text{by Rule 3}) \\ &\iff 0 \leq a^2 - 2ab + b^2 && (\text{by Rule 1}) \\ &\iff 0 \leq (a-b)^2. \end{aligned}$$

This final inequality is true, since the square of every real number is non-negative. It follows that the original inequality $ab \leq \left(\frac{a+b}{2}\right)^2$ is also true, for $a, b \in \mathbb{R}$.

If you examine the chain of equivalent statements in the solution to Worked Exercise D9, then you will see that we can replace the symbol \leq with $=$ throughout, without upsetting the equivalence of the statements. It follows that

$$ab = \left(\frac{a+b}{2}\right)^2 \text{ if and only if } a = b.$$

This tells us that the maximum area is obtained when the rectangle is a square.

In the next worked exercise you will see a proof of the following inequality:

$$\sqrt{a^2 + b^2} \leq a + b, \quad \text{for } a, b \geq 0.$$

This inequality has the geometric interpretation illustrated in Figure 11: the length of the hypotenuse of a right-angled triangle whose other sides are of lengths a and b is less than or equal to the sum of the lengths of those two sides.

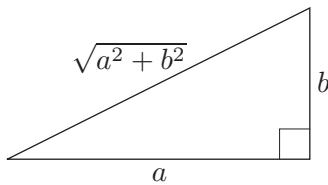


Figure 11 The lengths of the sides of a right-angled triangle

Worked Exercise D10

Prove that

$$\sqrt{a^2 + b^2} \leq a + b, \quad \text{for } a, b \geq 0.$$

Solution

 The first step is to remove the awkward square root by applying Rule 5, which we can do because $a, b \geq 0$. 

Rearranging the inequality, we obtain:

$$\begin{aligned}\sqrt{a^2 + b^2} \leq a + b &\iff a^2 + b^2 \leq (a + b)^2 \text{ (by Rule 5, with } p = 2\text{)} \\ &\iff a^2 + b^2 \leq a^2 + 2ab + b^2 \\ &\iff 0 \leq 2ab. \quad (\text{by Rule 1})\end{aligned}$$

This final inequality is certainly true, since $a, b \geq 0$. It follows that the original inequality $\sqrt{a^2 + b^2} \leq a + b$ is also true, for $a, b \geq 0$.

We can reformulate the inequality in Worked Exercise D10 in the following alternative way, which is sometimes useful in applications. If we write c in place of a^2 and d in place of b^2 , then the inequality becomes

$$\sqrt{c + d} \leq \sqrt{c} + \sqrt{d}, \quad \text{for } c, d \geq 0. \tag{2}$$

As an application, we use this new form of the inequality to prove that

$$|\sqrt{a} - \sqrt{b}| \leq \sqrt{|a - b|}, \quad \text{for } a, b \geq 0. \tag{3}$$

First note that interchanging the roles of a and b leaves inequality (3) unaltered, so it is sufficient to prove the inequality under the assumption that $a \geq b$. It follows from this assumption that $|a - b| = a - b$, and also that $\sqrt{a} \geq \sqrt{b}$, so that $|\sqrt{a} - \sqrt{b}| = \sqrt{a} - \sqrt{b}$. Hence

$$\begin{aligned}|\sqrt{a} - \sqrt{b}| \leq \sqrt{|a - b|} &\iff \sqrt{a} - \sqrt{b} \leq \sqrt{a - b} \\ &\iff \sqrt{a} \leq \sqrt{a - b} + \sqrt{b} \quad (\text{by Rule 2}).\end{aligned}$$

This final inequality is certainly true, since it can be obtained from inequality (2) simply by substituting $a - b$ in place of c , and b in place of d . It follows that inequality (3) is indeed true.

Exercise D12

Suppose that $a > 0$ and $a^2 > 2$. Prove that

$$\frac{1}{2} \left(a + \frac{2}{a} \right) < a.$$

3.3 Inequalities involving integers

In analysis we often need to prove inequalities involving an integer n . It is a common convention in mathematics that the symbol n is used to denote an integer, frequently a natural number. So, for example, the expression $n \geq 3$ means $n = 3, 4, 5, \dots$.

It is often possible to deal with inequalities involving integers by using the rules of rearrangement, just as we did in Subsection 3.2. The next worked exercise gives an example, and then there are two exercises for you to try.

Worked Exercise D11

Prove that

$$2n^2 \geq (n+1)^2, \quad \text{for } n \geq 3.$$

Solution

Rearranging this inequality into an equivalent form, we obtain

$$\begin{aligned} 2n^2 \geq (n+1)^2 &\iff 2n^2 - (n+1)^2 \geq 0 \quad (\text{by Rule 1}) \\ &\iff n^2 - 2n - 1 \geq 0 \\ &\iff (n-1)^2 - 2 \geq 0 \quad (\text{completing the square}) \\ &\iff (n-1)^2 \geq 2. \quad (\text{by Rule 1}) \end{aligned}$$

This final inequality is true for $n \geq 3$, so the original inequality $2n^2 \geq (n+1)^2$ is also true for $n \geq 3$.

An alternative solution to Worked Exercise D11 is the following:

$$\begin{aligned} 2n^2 \geq (n+1)^2 &\iff 2 \geq \left(\frac{n+1}{n}\right)^2 \quad (\text{by Rule 3}) \\ &\iff \sqrt{2} \geq 1 + \frac{1}{n} \quad (\text{by Rule 5, with } p = \tfrac{1}{2}). \end{aligned}$$

This final inequality certainly holds for $n \geq 3$, and so the original inequality holds too.

Exercise D13

Prove that

$$\frac{3n}{n^2 + 2} < 1, \quad \text{for } n > 2.$$

Exercise D14

Prove that

$$2n^3 \geq (n+1)^3, \quad \text{for } n \geq 4.$$

3.4 The Binomial Theorem

We often use a result known as the Binomial Theorem to prove inequalities involving integers. The Binomial Theorem gives us a general formula for the expansion of $(a + b)^n$, where $a, b \in \mathbb{R}$ and $n \in \mathbb{N}$. You will already be familiar with the special case when $n = 2$:

$$(a + b)^2 = a^2 + 2ab + b^2.$$

The statement of the Binomial Theorem uses the notation in the box below. Remember that the product $n \times (n - 1) \times \cdots \times 2 \times 1$ of the first n positive integers is denoted by the symbol $n!$, which is read as ‘ n factorial’ or ‘factorial n ’.

Definition

For any non-negative integers n and k with $k \leq n$,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

This expression is called a **binomial coefficient**. It is the number of combinations of n objects taken k at a time.

For example, if we have a set with three elements, and we want to know the number of different subsets with two elements, this is given by

$$\binom{3}{2} = \frac{3!}{2!1!} = 3.$$

Indeed, if we consider the set $\{a, b, c\}$, then the possible subsets with two elements are

$$\{a, b\}, \{b, c\} \text{ and } \{a, c\}.$$

Remarks

1. You may previously have met the alternative notation nC_k for binomial coefficients, instead of $\binom{n}{k}$, where the ‘ C ’ stands for ‘combination’. Both notations are in common use, and are sometimes read as ‘ n choose k ’.
2. We adopt the usual convention that $0! = 1$, so that

$$\binom{n}{0} = \binom{n}{n} = 1.$$

We can now state the Binomial Theorem.

Theorem D1 Binomial Theorem

If $a, b \in \mathbb{R}$ and $n \in \mathbb{N}$, then

$$\begin{aligned}(a+b)^n &= \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \\ &= a^n + na^{n-1}b + \frac{n(n-1)}{2!} a^{n-2}b^2 + \cdots + b^n.\end{aligned}$$

In the important special case where $a = 1$ and $b = x \in \mathbb{R}$, we have

$$\begin{aligned}(1+x)^n &= \sum_{k=0}^n \binom{n}{k} x^k \\ &= 1 + nx + \frac{n(n-1)}{2!} x^2 + \cdots + x^n.\end{aligned}$$

Proof We have

$$(a+b)^n = \underbrace{(a+b) \times (a+b) \times \cdots \times (a+b)}_{n \text{ times}}.$$

When this product is multiplied out, we find that each term of the form $a^{n-k}b^k$ arises by choosing the variable a from $n - k$ of the brackets and the variable b from the remaining k brackets. Thus the coefficient of $a^{n-k}b^k$ is equal to the number of ways of choosing a subset of $n - k$ brackets (or, equivalently, a subset of k brackets) from the set of n brackets, and this is precisely $\binom{n}{k}$, as required. ■

A striking mathematical pattern appears when we expand expressions of the form $(a+b)^n$ for $n = 1, 2, \dots$:

$$\begin{aligned}(a+b)^1 &= a^1 + b^1, \\ (a+b)^2 &= a^2 + 2ab + b^2, \\ (a+b)^3 &= (a+b)(a^2 + 2ab + b^2) = a^3 + 3a^2b + 3ab^2 + b^3, \\ (a+b)^4 &= (a+b)(a^3 + 3a^2b + 3ab^2 + b^3) = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4,\end{aligned}$$

and so on.

The coefficients that appear in these expansions can be arranged as a triangular table, known as *Pascal's triangle*. The entries on the left- and right-hand edges of the triangle are all 1s, and the remaining entries can be generated by using the rule that each entry is the sum of the two nearest entries in the row above. The 1 at the top corresponds to $n = 0$ since we have $(a + b)^0 = 1$.

$(a + b)^0$				1		
$(a + b)^1$			1	1		
$(a + b)^2$		1	2	1		
$(a + b)^3$	1	3	3	1		
$(a + b)^4$	1	4	6	4	1	
$(a + b)^5$	1	5	10	10	5	1
⋮		⋮				

Pascal's triangle is named after the French mathematician and philosopher Blaise Pascal (1623–1662). He was far from the first person to study this array of numbers, but his work on it in his *Traité du Triangle Arithmétique* was influential. Research on binomial coefficients was also carried out at about the same time by John Wallis (1616–1703) and then by Isaac Newton (1642–1727), who discovered that the Binomial Theorem can be generalised to negative and fractional powers.

Pascal's triangle had been studied centuries earlier by the Chinese mathematician Yang Hui (1238–1298) and the Persian astronomer and poet Omar Khayyam; in China, Pascal's triangle is known as the Yang Hui triangle.



Blaise Pascal

We now use the Binomial Theorem to prove some inequalities involving integers.

Worked Exercise D12

Prove that

$$2^n \geq 1 + n, \quad \text{for } n \geq 1.$$

Solution

By the Binomial Theorem, for $n \geq 1$ and $x \in \mathbb{R}$, we have

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \cdots + x^n.$$

 In the case $x \geq 0$, all the terms on the right are non-negative, so we can decrease the sum by omitting all but the first two terms. 

So

$$(1+x)^n \geq 1 + nx, \quad \text{for } x \geq 0.$$

If we now substitute $x = 1$ in this inequality, we obtain

$$2^n \geq 1 + n, \quad \text{for } n \geq 1.$$

This is the required result.

Worked Exercise D13

Prove that

$$2^{1/n} \leq 1 + \frac{1}{n}, \quad \text{for } n \geq 1.$$

Solution

 We start by rearranging the required result into an equivalent form. 

$$2^{1/n} \leq 1 + \frac{1}{n} \iff 2 \leq \left(1 + \frac{1}{n}\right)^n \quad (\text{by Rule 5, with } p = n)$$

 The bracket on the right can now be expanded using the special case of the Binomial Theorem and, as in the last worked exercise, we can then reduce the sum by omitting all but the first two terms. 

Applying the Binomial Theorem with $x = 1/n$, we get

$$\begin{aligned} \left(1 + \frac{1}{n}\right)^n &= 1 + n\left(\frac{1}{n}\right) + \frac{n(n-1)}{2!}\left(\frac{1}{n}\right)^2 + \cdots + \left(\frac{1}{n}\right)^n \\ &\geq 1 + 1 = 2. \end{aligned}$$

Thus the inequality $2 \leq \left(1 + \frac{1}{n}\right)^n$ is true for $n \geq 1$, so it follows that the original inequality

$$2^{1/n} \leq 1 + \frac{1}{n}, \quad \text{for } n \geq 1,$$

is also true, as required.

Exercise D15

Use the Binomial Theorem to prove that

$$\left(1 + \frac{1}{n}\right)^n \geq \frac{5}{2} - \frac{1}{2n}, \quad \text{for } n \geq 1.$$

Hint: Consider the first three terms in the binomial expansion.

3.5 Mathematical induction and Bernoulli's Inequality

Another useful tool for proving inequalities involving integers is the Principle of Mathematical Induction, which you met in Unit A3.

Principle of Mathematical Induction

To prove that a statement $P(n)$ is true for $n = 1, 2, \dots$:

1. show that $P(1)$ is true
2. show that the implication $P(k) \implies P(k+1)$ is true for $k = 1, 2, \dots$

Recall that the Principle can be adapted to prove that a statement $P(n)$ is true for all integers n greater than or equal to some given integer other than 1.

Worked Exercise D14

Prove that

$$2^n \geq n^2, \quad \text{for } n \geq 4.$$

Solution

Let $P(n)$ be the statement $2^n \geq n^2$.

We want to prove that $P(n)$ is true for $n \geq 4$, so we start with $P(4)$.

$P(4)$ is true since $2^4 = 16 = 4^2$.

Now let $k \geq 4$ and assume that $P(k)$ is true; that is,

$$2^k \geq k^2.$$

We wish to deduce that $P(k+1)$ is true; that is,

$$2^{k+1} \geq (k+1)^2.$$

Multiplying the inequality $2^k \geq k^2$ by 2 (using Rule 3) we get

$$2^{k+1} \geq 2k^2,$$

so it is sufficient for our purposes to prove that

$$2k^2 \geq (k+1)^2, \quad \text{for } k \geq 4.$$

Now,

$$\begin{aligned} 2k^2 \geq (k+1)^2 &\iff 2k^2 \geq k^2 + 2k + 1 \\ &\iff k^2 - 2k - 1 \geq 0 \quad (\text{by Rule 1}) \\ &\iff (k-1)^2 - 2 \geq 0 \quad (\text{completing the square}). \end{aligned}$$

This last inequality certainly holds for $k \geq 4$, so we have shown that

$$2^k \geq k^2 \implies 2^{k+1} \geq (k+1)^2, \quad \text{for } k \geq 4,$$

which was our aim.

Thus

$$P(k) \implies P(k+1), \quad \text{for } k \geq 4.$$

Hence, by mathematical induction, $P(n)$ is true, for $n \geq 4$.

Exercise D16

Prove that

$$2^n \geq n^3, \quad \text{for } n \geq 10.$$

Hint: You may find it helpful to use the solution to Exercise D14.

The next inequality, called *Bernoulli's Inequality*, will be used regularly in later units. We will prove the result using mathematical induction.



Bernoulli's Inequality is named after the Swiss mathematician Jacob Bernoulli (1654–1705) who published it in his *Positiones Arithmeticae de Seriebus Infinitis* (1689), using it several times. However, it is actually due to the Walloon mathematician René-François de Sluse (1622–1685), who published it in the second edition of his *Mesolabum* (1668).

Jacob Bernoulli was the first of the remarkable Bernoulli family who, over the course of three generations, produced eight gifted mathematicians. Jacob is best known for his *Ars Conjectandi*, a pioneering book on the theory of probability which was published posthumously in 1713 by his nephew, and for his founding work on the calculus of variations.

Jacob Bernoulli

René-François de Sluse's position in the church prevented him from visiting other mathematicians, but he corresponded with many mathematicians of the day, including Blaise Pascal.



Theorem D2 Bernoulli's Inequality

If $x \in \mathbb{R}$ and $n \in \mathbb{N}$, then

$$(1+x)^n \geq 1+nx, \quad \text{when } x \geq -1.$$

René-François de Sluse

Remarks

- In the solution to Worked Exercise D12 you saw that $(1+x)^n \geq 1+nx$ for $x \geq 0$ and $n \in \mathbb{N}$. Bernoulli's Inequality asserts that the same result holds under the *weaker* assumption that $x \geq -1$. Here, by a *weaker* assumption we mean an assumption that is less restrictive. Correspondingly, we say that Bernoulli's Inequality is a *stronger* result than that proved in the solution of Worked Exercise D12, because it applies more widely.
- If we sketch the graphs of $y = (1+x)^n$ and $y = 1+nx$, it certainly *looks* as though the first graph always lies above the second, so long as $x \geq -1$; see Figure 12. This *suggests* that Bernoulli's Inequality should hold for $x \geq -1$, but of course we need to prove it.

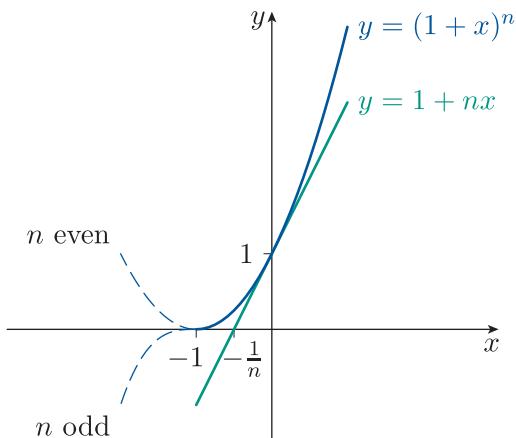


Figure 12 The graphs of $y = (1+x)^n$ and $y = 1+nx$

Proof of Theorem D2 Let $x \geq -1$ and let $P(n)$ be the statement $(1+x)^n \geq 1+nx$.

$P(1)$ is true since $(1+x)^1 = 1+x$.

Now let $k \geq 1$ and assume that $P(k)$ is true; that is,

$$(1+x)^k \geq 1+kx.$$

We wish to deduce that $P(k+1)$ is true; that is,

$$(1+x)^{k+1} \geq 1+(k+1)x.$$

Multiplying the inequality $(1+x)^k \geq 1+kx$ by the quantity $(1+x)$, which is non-negative because $x \geq -1$, we get

$$\begin{aligned}(1+x)^{k+1} &\geq (1+x)(1+kx) \\ &= 1+(k+1)x+kx^2 \\ &\geq 1+(k+1)x,\end{aligned}$$

since the term kx^2 is positive. Thus we have shown that

$$P(k) \implies P(k+1), \quad \text{for } k \geq 1.$$

Hence, by mathematical induction, $P(n)$ is true, for $n \geq 1$. ■

In Worked Exercise D13 you saw that

$$2^{1/n} \leq 1 + \frac{1}{n}, \quad \text{for } n \geq 1.$$

In the next worked exercise you will see how Bernoulli's Inequality can be used to prove another inequality involving $2^{1/n}$.

Worked Exercise D15

By applying Bernoulli's Inequality with $x = -1/(2n)$, prove that

$$2^{1/n} \geq 1 + \frac{1}{2n-1}, \quad \text{for } n \geq 1.$$

Solution

 We can apply Bernoulli's Inequality because $-1/(2n) \geq -1$ for $n \geq 1$. 

Substituting $x = -1/(2n)$ into Bernoulli's Inequality, we obtain

$$\begin{aligned}\left(1 - \frac{1}{2n}\right)^n &\geq 1 + n\left(-\frac{1}{2n}\right) \\ &= \frac{1}{2}.\end{aligned}$$

If we then take the n th root of both sides of this inequality (which is permissible, by Rule 5), we obtain

$$1 - \frac{1}{2n} \geq \frac{1}{2^{1/n}},$$

that is,

$$\frac{2n-1}{2n} \geq \frac{1}{2^{1/n}}$$

hence, by Rule 4,

$$2^{1/n} \geq \frac{2n}{2n-1}.$$

We now write $2n = (2n-1) + 1$ to get the right-hand side into the required form, and this completes the proof.

Hence,

$$2^{1/n} \geq \frac{(2n-1)+1}{2n-1} = 1 + \frac{1}{2n-1}, \quad \text{for } n \geq 1.$$

Combining the results of Worked Exercises D13 and D15, we have shown that

$$1 + \frac{1}{2n-1} \leq 2^{1/n} \leq 1 + \frac{1}{n}, \quad \text{for } n \geq 1.$$

Exercise D17

By applying Bernoulli's Inequality, first with $x = -3/(4n)$, and then with $x = 3/n$, prove that

$$1 + \frac{3}{4n-3} \leq 4^{1/n} \leq 1 + \frac{3}{n}, \quad \text{for } n \geq 1.$$

4 Least upper bounds

In this section you will meet the idea of upper and lower bounds of a set and then see how to identify the least upper bound and the greatest lower bound of a set. You will also learn about the Least Upper Bound Property of \mathbb{R} .

4.1 Upper bounds and lower bounds

Any finite set of real numbers has a greatest element (and a least element), but this property does not necessarily hold for sets with infinitely many elements. For example, neither of the sets $\mathbb{N} = \{1, 2, 3, \dots\}$ and $[0, 2)$ has a greatest element. However, the set $[0, 2)$ is *bounded above* by 2, since all points of $[0, 2)$ are less than 2.

Definitions

A set $E \subseteq \mathbb{R}$ is **bounded above** if there is a real number M , called an **upper bound** of E , such that

$$x \leq M, \quad \text{for all } x \in E.$$

If the upper bound M belongs to E , then M is called the **maximum element** of E , and is denoted by $\max E$.

Geometrically, the set E is bounded above by M if no point of E lies to the right of M on the real line.

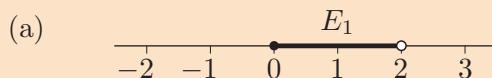
For example, if $E = [0, 2)$, then the numbers 2, 3, 3.5 and 157.1 are all upper bounds of E , whereas the numbers 1.995, 1.5, 0 and -157.1 are not upper bounds of E . Although it may seem obvious that $[0, 2)$ has no maximum element, you may find it difficult to write down a proof of this fact. The next worked exercise demonstrates how to do this.

Worked Exercise D16

Sketch the following sets on the real line, and determine which are bounded above and which have a maximum element.

- (a) $E_1 = [0, 2)$
- (b) $E_2 = \{1/n : n = 1, 2, \dots\}$
- (c) $E_3 = \mathbb{N}$

Solution



The set E_1 is bounded above. For example, $M = 2$ is an upper bound of E_1 , since

$$x \leq 2, \quad \text{for all } x \in E_1.$$

However, E_1 has no maximum element, as we now show.

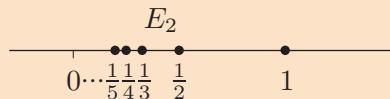
The number 2 is not a maximum element, since $2 \notin E_1$. We choose a general element in E_1 and show that there is another element in E_1 which is greater.

For each x in E_1 we have $x < 2$, so there is a real number y such that

$$x < y < 2,$$

by the Density Property of \mathbb{R} . Hence $y \in E_1$, so x is not a maximum element of E_1 . This shows that no element of E_1 is a maximum element of E_1 .

(b)



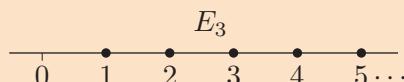
The set E_2 is bounded above. For example, $M = 1$ is an upper bound of E_2 , since

$$\frac{1}{n} \leq 1, \quad \text{for } n = 1, 2, \dots$$

Also, since $1/n = 1$ when $n = 1$, we have $1 \in E_2$ and so

$$\max E_2 = 1.$$

(c)



The set E_3 is not bounded above. For each real number M , there is a positive integer n such that $n > M$, by the Archimedean Property of \mathbb{R} . Hence M is not an upper bound of E_3 .

Since E_3 is not bounded above, it has no maximum element.

Exercise D18

Sketch the following sets on the real line, and determine which are bounded above and which have a maximum element.

- (a) $E_1 = (-\infty, 1]$
- (b) $E_2 = \{1 - 1/n : n = 1, 2, \dots\}$
- (c) $E_3 = \{n^2 : n = 1, 2, \dots\}$

Lower bounds are defined in a similar way to upper bounds. For example, the interval $(0, 2)$ is bounded below by 0, since

$$0 \leq x, \quad \text{for all } x \in (0, 2).$$

However, 0 does not belong to $(0, 2)$, so 0 is not a minimum element of $(0, 2)$. In fact, $(0, 2)$ has no minimum element.

Definitions

A set $E \subseteq \mathbb{R}$ is **bounded below** if there is a real number m , called a **lower bound** of E , such that

$$m \leq x, \quad \text{for all } x \in E.$$

If the lower bound m belongs to E , then m is called the **minimum element** of E , and is denoted by $\min E$.

Geometrically, the set E is bounded below by m if no point of E lies to the left of m on the real line.

Exercise D19

Determine which of the following sets are bounded below and which have a minimum element. (The sketches you made in Exercise D18 may help.)

- (a) $E_1 = (-\infty, 1]$
- (b) $E_2 = \{1 - 1/n : n = 1, 2, \dots\}$
- (c) $E_3 = \{n^2 : n = 1, 2, \dots\}$

The following terminology is also useful.

Definitions

A set $E \subseteq \mathbb{R}$ is **bounded** if E is bounded above and bounded below; the set E is **unbounded** if it is not bounded.

For example, of the sets you met in Exercises D18 and D19, the set $E_2 = \{1 - 1/n : n = 1, 2, \dots\}$ is bounded, but $E_1 = (-\infty, 1]$ and $E_3 = \{n^2 : n = 1, 2, \dots\}$ are unbounded.

4.2 Least upper bounds and greatest lower bounds

We have seen that the set $[0, 2)$ has no maximum element. However, $[0, 2)$ has many upper bounds, for example 2, 3, 3.5 and 157.1. Among all these upper bounds, the number 2 is the *least* upper bound, because any number less than 2 is not an upper bound of $[0, 2)$, as illustrated in Figure 13.

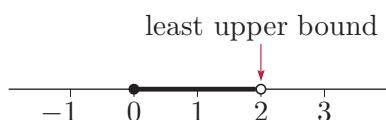


Figure 13 The least upper bound of $[0, 2)$

The least upper bound of a set is also called the *supremum* of a set. This comes from the Latin word *supremus* meaning ‘highest’.

Definition

A real number M is the **least upper bound**, or **supremum**, of a set $E \subseteq \mathbb{R}$ if

1. M is an upper bound of E
2. each number $M' < M$ is not an upper bound of E .

In this case, we write $M = \sup E$.

If E has a maximum element, $\max E$, then $\sup E = \max E$. For example, the closed interval $[0, 2]$ has maximum element 2, so it has least upper bound 2.

If a set does not have a maximum element but is bounded above, then we may be able to guess the value of its least upper bound. As in the example $E = [0, 2)$, there may be an obvious ‘missing point’ at the upper end of the set. We now see how to *prove* that our guess is correct.

Worked Exercise D17

Prove that the least upper bound of $[0, 2)$ is 2.

Solution

First check that 2 is *an* upper bound.

We know that $M = 2$ is an upper bound of $[0, 2)$ because

$$x \leq 2, \quad \text{for all } x \in [0, 2).$$

To show that 2 is the *least* upper bound, we must prove that each number $M' < 2$ is *not* an upper bound of $[0, 2)$.

Suppose that $M' < 2$. We must find an element x in $[0, 2)$ which is greater than M' . By the Density Property, there is a real number x such that

$$M' < x < 2 \quad \text{and} \quad x \geq 0.$$

Thus $x \in [0, 2)$ and $x > M'$, which shows that M' is not an upper bound of $[0, 2)$. Hence $M = 2$ is the least upper bound of $[0, 2)$.

The solution to Worked Exercise D17 illustrates the following strategy for determining the least upper bound of a set, if there is one.

Strategy D1

To show that M is the least upper bound (supremum) of a subset E of \mathbb{R} , check that:

1. $x \leq M$, for all $x \in E$
2. if $M' < M$, then there is some $x \in E$ such that $x > M'$.

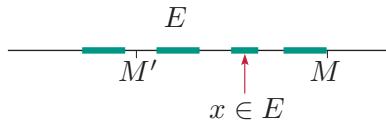


Figure 14 The points and sets in Strategy D1

The points and sets involved in Strategy D1 are illustrated in Figure 14, with the set E being indicated by bold green lines. If M is an upper bound of a set E and $M \in E$, then steps 1 and 2 of the strategy are automatically satisfied, so $M = \sup E = \max E$.

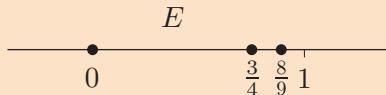
Worked Exercise D18

Determine the least upper bound of the set

$$E = \{1 - 1/n^2 : n = 1, 2, \dots\}.$$

Solution

>We first guess the least upper bound of E . A sketch may help.



We guess from the sketch that the least upper bound of E is $M = 1$. We now use Strategy D1.

We have that 1 is an upper bound of E , since

$$1 - \frac{1}{n^2} \leq 1, \quad \text{for } n = 1, 2, \dots$$

We now need to show that, if $M' < 1$, then M' is not an upper bound of E ; that is, there is some natural number n such that

$$1 - \frac{1}{n^2} > M'.$$

We show this by rearranging the inequality into an equivalent form with just n on one side.

Suppose that $0 < M' < 1$. We have

$$\begin{aligned} 1 - \frac{1}{n^2} > M' &\iff 1 - M' > \frac{1}{n^2} \\ &\iff \frac{1}{1 - M'} < n^2 \\ &\iff \sqrt{\frac{1}{1 - M'}} < n \quad (\text{since } (1 - M') > 0 \text{ and } n > 0). \end{aligned}$$

We can certainly choose n so that this final inequality holds, by the Archimedean Property of \mathbb{R} . Hence, for this value of n , we have

$$1 - \frac{1}{n^2} > M',$$

so M' is not an upper bound of E .

It follows that 1 is the least upper bound of E .

Exercise D20

Determine the least upper bound, if it exists, for each of the following sets. (You will find it helpful to refer to your solution to Exercise D18.)

- (a) $E_1 = (-\infty, 1]$
- (b) $E_2 = \{1 - 1/n : n = 1, 2, \dots\}$
- (c) $E_3 = \{n^2 : n = 1, 2, \dots\}$

The *greatest lower bound* or *infimum* of a set is defined in a similar way to the least upper bound. The word infimum comes from the Latin word *infimus* meaning ‘least’.

Definition

A real number m is the **greatest lower bound**, or **infimum**, of a set $E \subseteq \mathbb{R}$ if

1. m is a lower bound of E
2. each number $m' > m$ is not a lower bound of E .

In this case, we write $m = \inf E$.

If E has a minimum element, then $\inf E = \min E$. For example, the closed interval $[0, 2]$ has minimum element 0, so it has greatest lower bound 0.

The following strategy for proving that a number is the greatest lower bound of a set is similar to Strategy D1.

Strategy D2

To show that m is the greatest lower bound (infimum) of a subset E of \mathbb{R} , check that:

1. $x \geq m$, for all $x \in E$
2. if $m' > m$, then there is some $x \in E$ such that $x < m'$.

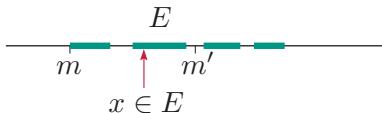


Figure 15 The sets and points in Strategy D2

Exercise D21

Determine the greatest lower bound, if it exists, for each of the following sets.

- (a) $E_1 = (1, 5]$ (b) $E_2 = \{1/n^2 : n = 1, 2, \dots\}$

4.3 Least Upper Bound Property

In the exercises and worked exercises in the previous subsection it was straightforward to guess the values of $\sup E$ and $\inf E$. Sometimes, however, this is not the case. For example, if

$$E = \{(1 + 1/n)^n : n = 1, 2, \dots\} = \left\{ \left(\frac{2}{1}\right)^1, \left(\frac{3}{2}\right)^2, \left(\frac{4}{3}\right)^3, \dots \right\},$$

then it is not very easy to guess the value of the least upper bound of E . It turns out that

$$\sup E = e = 2.71828\dots$$

In cases like this it can be reassuring to know that $\sup E$ does exist, even though it may be difficult to find.

The existence of $\sup E$ is guaranteed by the following fundamental result, which is the basis for many other results in analysis. (This is an example of an *existence theorem*, that is, a theorem that asserts that a mathematical object, such as a real number with a certain property, must exist, but does not tell you what it is.)

Least Upper Bound Property of \mathbb{R}

Let E be a non-empty subset of \mathbb{R} . If E is bounded above, then E has a least upper bound.

This property of \mathbb{R} is very plausible geometrically. If the set E lies entirely to the left of some number M , then you can imagine decreasing the value of M steadily until any further decrease causes M to be less than some point of E . At this point, $\sup E$ has been reached, as shown in Figure 16, where the set E is indicated by bold green lines.

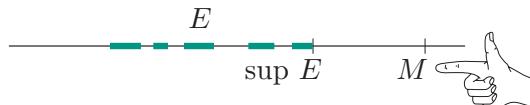


Figure 16 The Least Upper Bound Property of \mathbb{R}

The Least Upper Bound Property of \mathbb{R} can be used to show that \mathbb{R} does indeed include non-recurring decimals which represent irrational numbers such as $\sqrt{2}$, as was claimed in Section 1. It can also be used to define the arithmetic operations of addition and multiplication with such decimals. We discuss this further in Section 5.

As you would expect, there is a corresponding property for lower bounds.

Greatest Lower Bound Property of \mathbb{R}

Let E be a non-empty subset of \mathbb{R} . If E is bounded below, then E has a greatest lower bound.

We now give a proof of the Least Upper Bound Property in the case when the set E contains at least one positive number. The proof in the general case can be reduced to this special case, but we do not give the details. Try to follow the proof if you are interested, but you can omit it if you are short of time.

Proof of the Least Upper Bound Property Let E be a subset of \mathbb{R} that is bounded above and contains at least one positive number.

➊ We first choose a suitable candidate for the least upper bound, and then use Strategy D1. ➋

We apply the following procedure to give us the successive digits in a decimal $a_0.a_1a_2\dots$, which we then prove to be the least upper bound of E .

We choose in succession:

the greatest integer a_0 such that a_0 is not an upper bound of E

the greatest digit a_1 such that $a_0.a_1$ is not an upper bound of E

the greatest digit a_2 such that $a_0.a_1a_2$ is not an upper bound of E

⋮

the greatest digit a_n such that $a_0.a_1\dots a_n$ is not an upper bound of E

⋮

Thus at the n th stage, the selected digit a_n has the properties that

$$a_0.a_1a_2\dots a_n \text{ is not an upper bound of } E \quad (4)$$

$$a_0.a_1a_2\dots a_n + \frac{1}{10^n} \text{ is an upper bound of } E. \quad (5)$$

 Since E contains at least one positive number, the numbers in (4) and (5) are positive rationals. 

We now use Strategy D1 to prove that the least upper bound of E is the non-terminating decimal

$$a = a_0.a_1a_2\dots$$

First we have to show that a is an upper bound of E ; that is, if $x \in E$, then $x \leq a$. To do this, we prove the equivalent contrapositive statement: if $x > a$, then $x \notin E$.

 Note that our procedure for choosing a always produces a non-terminating decimal. For example, it would give $a = 1.49999\dots$, rather than $a = 1.5$. For the purpose of making comparisons, we therefore represent x in the same way. 

Let $x > a$, and represent x as a non-terminating decimal $x = x_0.x_1x_2\dots$

Since $x > a$, there is a least integer n such that $x_n > a_n$, which means that $x_n \geq a_n + 1$. Thus for this value of n we have

$$x_0.x_1x_2\dots x_n \geq a_0.a_1a_2\dots a_n + \frac{1}{10^n} > a_0.a_1a_2\dots,$$

so $x_0.x_1x_2\dots x_n$ is an upper bound of E , by statement (5). Since x is non-terminating, $x > x_0.x_1x_2\dots x_n$, so $x \notin E$, as required.

To complete the proof we have to show that if $a' < a$, then a' is not an upper bound of E . Let $a' < a$. Then there is an integer n such that

$$a' < a_0.a_1a_2\dots a_n,$$

so a' is not an upper bound of E , by statement (4).

Thus we have proved that a is the least upper bound of E . 

5 Manipulating real numbers

This section is intended for reading only. There are no exercises on this section.

5.1 Arithmetic with real numbers

At the end of Section 1 we discussed the non-recurring decimals representing the irrationals $\sqrt{2}$ and π , which begin

$$\sqrt{2} = 1.414\ 213\ 56\dots \quad \text{and} \quad \pi = 3.141\ 592\ 65\dots,$$

and asked whether it is possible to add and multiply these numbers to obtain another real number. We now explain how this can be done using the Least Upper Bound Property of \mathbb{R} .

A natural way to obtain a sequence of approximations to the sum $\sqrt{2} + \pi$ is to truncate each of the above decimals and then form the sums of these truncated decimals. If each of the decimals is truncated at the same decimal place, then we obtain the following sequences of approximations, which are increasing.

$\sqrt{2}$	π	$\sqrt{2} + \pi$
1	3	4
1.4	3.1	4.5
1.41	3.14	4.55
1.414	3.141	4.555
1.4142	3.1415	4.5557
:	:	:

Intuitively we expect that the sum $\sqrt{2} + \pi$ is greater than each of the numbers in the right-hand column, but ‘only just’. To accord with our intuition, therefore, we *define* the sum $\sqrt{2} + \pi$ to be the least upper bound of the set of numbers in the right-hand column; that is,

$$\sqrt{2} + \pi = \sup\{4, 4.5, 4.55, 4.555, 4.5557, \dots\}.$$

To be sure that this definition makes sense, we need to show that this set is bounded above. But all the truncations of $\sqrt{2}$ are less than 1.5 and all those of π are less than 4. Hence, all the sums in the right-hand column are less than $1.5 + 4 = 5.5$. So, by the Least Upper Bound Property of \mathbb{R} , the set of numbers in the right-hand column *does* have a least upper bound and we *can* define $\sqrt{2} + \pi$ this way.

This method can be used to define the sum of any pair of positive real numbers.

Let us check that this method of adding decimals gives the correct answer when we use it in a familiar case. Consider the simple calculation

$$\frac{1}{3} + \frac{2}{3} = 0.333\dots + 0.666\dots$$

Truncating each of these decimals and forming the sums, we obtain the set

$$\{0, 0.9, 0.99, 0.999, \dots\}.$$

The supremum of this set is $0.999\dots = 1$, which is the correct answer. (We are not suggesting that this is a practical method for adding rationals!)

We can define the product of any two positive real numbers in a similar way. For example, to define $\sqrt{2} \times \pi$ we can form the sequence of products of their truncations.

$\sqrt{2}$	π	$\sqrt{2} \times \pi$
1	3	3
1.4	3.1	4.34
1.41	3.14	4.4274
1.414	3.141	4.441374
1.4142	3.1415	4.4427093
:	:	:

As before, we define $\sqrt{2} \times \pi$ to be the least upper bound of the set of numbers in the right-hand column.

Similar ideas can be used to define the operations of subtraction and division, but we omit the details.

In this way we can define arithmetic with real numbers in terms of the familiar arithmetic with rationals by using the Least Upper Bound Property of \mathbb{R} . Moreover, it can be proved that these operations in \mathbb{R} satisfy all the usual properties of a field; you saw these listed in Subsection 1.5.

5.2 Existence of roots

Just as we usually take for granted the basic arithmetic operations with real numbers, so we usually assume that, given any positive real number a , there is a unique positive real number $b = \sqrt{a}$ such that $b^2 = a$. We now discuss the justification for this assumption.

First, here is a geometric justification. Given line segments of lengths 1 and a , we can construct a semicircle with diameter $a + 1$ as shown in Figure 17.

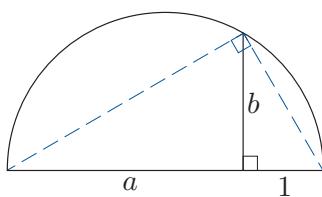


Figure 17 A semicircle with diameter $a + 1$

Using similar triangles, we find that

$$\frac{a}{b} = \frac{b}{1}, \quad \text{so} \quad b^2 = a.$$

This shows that there should be a positive real number b such that $b^2 = a$, in order that the length of the vertical line segment in this figure can be described exactly. But does $b = \sqrt{a}$ always exist *exactly* as a real number? In fact it does, and a more general result is true.

Theorem D3

For each positive real number a and each integer $n > 1$, there is a unique positive real number b such that

$$b^n = a.$$

We call this positive number b the **n th root** of a , and write $b = \sqrt[n]{a}$. We also define $\sqrt[0]{0} = 0$, since $0^n = 0$.

In the special case $a = 2$ and $n = 2$, Theorem D3 asserts the existence of a positive real number b such that

$$b^2 = 2.$$

Here is a direct proof of Theorem D3 in this special case. (A proof of the general case is given in Section 4 of Unit D4.) We choose the numbers $1, 1.4, 1.41, 1.414, \dots$ to satisfy the inequalities:

$$\left. \begin{array}{l} 1^2 < 2 < 2^2 \\ (1.4)^2 < 2 < (1.5)^2 \\ (1.41)^2 < 2 < (1.42)^2 \\ (1.414)^2 < 2 < (1.415)^2 \\ \vdots \end{array} \right\} \quad (6)$$

This process gives an infinite decimal $b = 1.414\dots$ and we claim that

$$b^2 = (1.414\dots)^2 = 2.$$

This can be proved using our method of multiplying decimals.

b	b	b^2
1	1	1
1.4	1.4	1.96
1.41	1.41	1.9881
1.414	1.414	1.999396
\vdots	\vdots	\vdots

We have to prove that the least upper bound of the set E of numbers in the right-hand column is 2. In other words,

$$\sup E = \sup\{1, 1.4^2, 1.41^2, 1.414^2, \dots\} = 2.$$

To do this, we use Strategy D1.

First we check that $M = 2$ is an upper bound of E . This follows from the left-hand inequalities in (6).

Next we check that if $M' < 2$, then there is a number in E which is greater than M' . To prove this, we put

$$x_0 = 1, \quad x_1 = 1.4, \quad x_2 = 1.41, \quad x_3 = 1.414, \quad \dots$$

By the right-hand inequalities in (6) we have, for $n = 0, 1, 2, \dots$,

$$2 < \left(x_n + \frac{1}{10^n} \right)^2 = x_n^2 + \frac{2x_n}{10^n} + \left(\frac{1}{10^n} \right)^2,$$

so

$$x_n^2 > 2 - \frac{1}{10^n} \left(2x_n + \frac{1}{10^n} \right).$$

Since $x_n < 2$, we have

$$2x_n + \frac{1}{10^n} < 2 \times 2 + 1 = 5$$

and so

$$x_n^2 > 2 - \frac{5}{10^n} = 1.\underbrace{99\dots95}_{n \text{ digits}}.$$

Thus if $M' < 2$, then we *can* choose n so large that $x_n^2 > M'$. This proves that the least upper bound of the set E is 2, so

$$b^2 = (1.414\dots)^2 = 2,$$

as claimed. Thus $b = 1.414\dots$ is a decimal representation of $\sqrt{2}$.

5.3 Powers

Having discussed n th roots, we are now in a position to define the expression a^x , where a is positive and x is a rational power (or exponent).

Definition

If $a > 0$, $m \in \mathbb{Z}$ and $n \in \mathbb{N}$, then

$$a^{m/n} = (\sqrt[n]{a})^m$$

or, equivalently,

$$a^{m/n} = \sqrt[n]{a^m}.$$

For example, for $a > 0$ with $m = 1$ we have $a^{1/n} = \sqrt[n]{a}$, and with $m = 2$ and $n = 3$ we have $a^{2/3} = (\sqrt[3]{a})^2$.

This notation is particularly useful because rational powers satisfy the following laws, whose proofs depend on Theorem D3.

Index Laws

If $a, b > 0$ and $x, y \in \mathbb{Q}$, then

$$a^x b^x = (ab)^x, \quad a^x a^y = a^{x+y} \quad \text{and} \quad (a^x)^y = a^{xy}.$$

Remarks

1. If x and y are *integers*, then these laws also hold for all non-zero real numbers a and b , not just positive ones. However, if x and y are not integers, then we must have $a > 0$ and $b > 0$. For example, $(-1)^{1/2}$ is not defined as a real number.
2. Each positive number has *two* n th roots when n is even. For example, $2^2 = (-2)^2 = 4$. We adopt the convention that, for $a > 0$, $\sqrt[n]{a}$ and $a^{1/n}$ always denote the *positive* n th root of a . If we wish to refer to both roots (for example, when solving equations), then we write $\pm\sqrt[n]{a}$.

We conclude this section by briefly discussing the meaning of a^x when $a > 0$ and x is an arbitrary real number. We have defined a^x when x is rational, but the same definition does not work if x is irrational. However, it is common practice to write expressions such as

$$\sqrt{2}^{\sqrt{2}}$$

and even to apply the Index Laws to give equalities such as

$$\left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}} = \sqrt{2}^{\sqrt{2} \times \sqrt{2}} = (\sqrt{2})^2 = 2.$$

Such manipulations *can* be justified, and by the end of this book you will have seen one way to do this. Moreover, the justification uses several key ideas from the book, including convergence of sequences, convergence of series and continuity.

Summary

In this unit you have met many of the basic ideas and techniques you will need as you continue your study of analysis.

You have seen how the real numbers and their arithmetic operations can be precisely defined using infinite decimals. You have learned how to manipulate, solve and prove inequalities, and met the Triangle Inequality and Bernoulli's Inequality, both of which are widely used in analysis. You have studied upper and lower bounds of sets of real numbers, and discovered how to show whether a given number is a least upper bound or a greatest lower bound of such a set.

Finally, you have seen that whilst both the rationals and the reals are ordered fields, only the real numbers possess the Least Upper Bound Property, namely that every non-empty subset of \mathbb{R} that is bounded above has a least upper bound. This is the key property that distinguishes the reals from the rationals, and it underlies many of the ideas you will meet in future analysis units.

Learning outcomes

After working through this unit, you should be able to:

- explain the relationship between *rational numbers* and *recurring decimals*
- explain the term *irrational number*
- understand how the *real numbers* can be defined in terms of infinite decimals
- find a rational and an irrational number between any two distinct real numbers
- solve inequalities by rearranging them into simpler equivalent forms
- solve inequalities involving modulus signs
- use the Triangle Inequality
- use the Binomial Theorem, mathematical induction and Bernoulli's Inequality to prove inequalities which involve an integer n
- explain the terms *bounded above*, *bounded below* and *bounded*
- use the strategies for determining least upper bounds and greatest lower bounds
- state the Least Upper Bound Property of \mathbb{R} and the Greatest Lower Bound Property of \mathbb{R}
- explain how the Least Upper Bound Property is used to define arithmetic operations with real numbers
- explain the meaning of rational powers.

Solutions to exercises

Solution to Exercise D1

Since $45 \times 20 = 900$ and $17 \times 53 = 901$, we have $45/53 < 17/20$. Thus

$$-1 < -\frac{17}{20} < -\frac{45}{53} < 0 < \frac{45}{53} < \frac{17}{20} < 1.$$

Solution to Exercise D2

Let a, b be distinct rationals, where $a < b$.

Let $c = \frac{1}{2}(a + b)$; then c is rational, and

$$c - a = \frac{1}{2}(b - a) > 0,$$

$$b - c = \frac{1}{2}(b - a) > 0,$$

so $a < c < b$.

Solution to Exercise D3

We have $\frac{1}{7} = 0.\overline{142857}$

Solution to Exercise D4

(a) Let $x = 0.\overline{231}$.

Multiplying both sides by 10^3 , we obtain

$$1000x = 231.\overline{231} = 231 + x.$$

Hence

$$999x = 231, \quad \text{so} \quad x = \frac{231}{999} = \frac{77}{333}.$$

(b) Let $x = 0.\overline{81}$.

Multiplying both sides by 10^2 , we obtain

$$100x = 81.\overline{81} = 81 + x.$$

Hence

$$99x = 81, \quad \text{so} \quad x = \frac{81}{99} = \frac{9}{11}.$$

Thus

$$2.2\overline{81} = 2 + \frac{2}{10} + \frac{9}{110} = \frac{251}{110}.$$

Solution to Exercise D5

$\frac{17}{20} = 0.85$ and $\frac{45}{53} = 0.84\dots$, so $\frac{45}{53} < \frac{17}{20}$.

Solution to Exercise D6

$x = 0.34$ and $y = 0.340\,010\,010\,001\dots$,

where $010\,010\,001\dots$ is a non-recurring tail.

Solution to Exercise D7

(a) Rule 2 with $c = 3$ gives $x + 3 > 5$.

(b) Rule 1 followed by Rule 3 with $c = -1$ gives $2 - x < 0$.

(c) Rule 3 with $c = 5$ followed by Rule 2 with $c = 2$ gives $5x + 2 > 12$.

(d) Part (c) followed by Rule 4 gives $1/(5x + 2) < 1/12$.

Solution to Exercise D8

(a) Note that this inequality cannot be solved by cross-multiplying, because $x^2 - 1$ can be positive, zero or negative, depending on the value of x .

Rearranging the inequality (using Rule 1 and Rule 3 with $c = -4$), we obtain

$$\begin{aligned} \frac{4x - x^2 - 7}{x^2 - 1} \geq 3 &\iff \frac{4x - x^2 - 7}{x^2 - 1} - 3 \geq 0 \\ &\iff \frac{4x - 4x^2 - 4}{x^2 - 1} \geq 0 \\ &\iff \frac{x^2 - x + 1}{x^2 - 1} \leq 0 \\ &\iff \frac{(x - \frac{1}{2})^2 + \frac{3}{4}}{x^2 - 1} \leq 0. \end{aligned}$$

Since $(x - \frac{1}{2})^2 + \frac{3}{4} > 0$, for all x , the inequality holds if and only if $x^2 - 1 = (x - 1)(x + 1) < 0$. (The fraction is undefined when $x^2 - 1 = 0$.)

Hence the solution set is

$$\left\{ x : \frac{4x - x^2 - 7}{x^2 - 1} \geq 3 \right\} = (-1, 1).$$

(b) Rearranging the inequality (using Rule 1), we obtain

$$\begin{aligned} 2x^2 \geq (x + 1)^2 &\iff 2x^2 \geq x^2 + 2x + 1 \\ &\iff x^2 - 2x - 1 \geq 0 \\ &\iff (x - 1)^2 - 2 \geq 0 \\ &\iff (x - 1)^2 \geq 2. \end{aligned}$$

Hence the solution set is

$$\begin{aligned} & \{x : 2x^2 \geq (x+1)^2\} \\ &= \{x : x-1 \leq -\sqrt{2}\} \cup \{x : x-1 \geq \sqrt{2}\} \\ &= (-\infty, 1-\sqrt{2}] \cup [1+\sqrt{2}, \infty). \end{aligned}$$

Solution to Exercise D9

The expression $\sqrt{2x^2 - 2}$ is defined, and non-negative, when $2x^2 - 2 \geq 0$, that is, for $x^2 \geq 1$. Thus $\sqrt{2x^2 - 2}$ is defined if x lies in $(-\infty, -1] \cup [1, \infty)$.

For $x \geq 1$, using Rule 5 with $p = 2$, we see that

$$\begin{aligned} \sqrt{2x^2 - 2} > x &\iff 2x^2 - 2 > x^2 \\ &\iff x^2 > 2. \end{aligned}$$

So the part of the solution set in $[1, \infty)$ is $(\sqrt{2}, \infty)$.

For $x \leq -1$,

$$\sqrt{2x^2 - 2} \geq 0 > x,$$

so the whole of $(-\infty, -1]$ lies in the solution set.

Hence the complete solution set is

$$\{x : \sqrt{2x^2 - 2} > x\} = (-\infty, -1] \cup (\sqrt{2}, \infty).$$

Solution to Exercise D10

(a) Rearranging the inequality (using Rule 6 and Rule 5 with $p = 2$), we obtain

$$\begin{aligned} |2x^2 - 13| < 5 &\iff -5 < 2x^2 - 13 < 5 \\ &\iff 8 < 2x^2 < 18 \\ &\iff 4 < x^2 < 9 \\ &\iff 4 < |x|^2 < 9 \\ &\iff 2 < |x| < 3. \end{aligned}$$

Hence the solution set is

$$\{x : |2x^2 - 13| < 5\} = (-3, -2) \cup (2, 3).$$

(b) Rearranging the inequality (using Rule 5 with $p = 2$), we have

$$\begin{aligned} |x-1| \leq 2|x+1| &\iff (x-1)^2 \leq 4(x+1)^2 \\ &\iff x^2 - 2x + 1 \leq 4x^2 + 8x + 4 \\ &\iff 0 \leq 3x^2 + 10x + 3 \\ &\iff 0 \leq (3x+1)(x+3). \end{aligned}$$

Hence the solution set is

$$\{x : |x-1| \leq 2|x+1|\} = (-\infty, -3] \cup [-\frac{1}{3}, \infty).$$

Solution to Exercise D11

(a) Suppose that $|a| \leq \frac{1}{2}$.

The Triangle Inequality gives

$$\begin{aligned} |a+1| &\leq |a| + |1| \\ &\leq \frac{1}{2} + 1 \quad (\text{since } |a| \leq \frac{1}{2}) \\ &= \frac{3}{2}. \end{aligned}$$

Hence

$$|a| \leq \frac{1}{2} \implies |a+1| \leq \frac{3}{2}.$$

(b) Suppose that $|b| < \frac{1}{2}$.

The backwards form of the Triangle Inequality gives

$$\begin{aligned} |b^3 - 1| &\geq ||b^3| - |1|| \\ &= ||b|^3 - 1| \\ &\geq 1 - |b|^3. \end{aligned}$$

Now

$$|b| < \frac{1}{2} \implies |b|^3 < \frac{1}{8} \implies 1 - |b|^3 > \frac{7}{8},$$

so, from the previous chain of inequalities,

$$|b| < \frac{1}{2} \implies |b^3 - 1| > \frac{7}{8}.$$

Solution to Exercise D12

Rearranging the inequality (using Rule 2 and Rule 3), we obtain

$$\begin{aligned} \frac{1}{2} \left(a + \frac{2}{a} \right) < a &\iff \frac{1}{2}a + \frac{1}{a} < a \\ &\iff \frac{1}{a} < \frac{1}{2}a \\ &\iff 2 < a^2. \end{aligned}$$

Since the final inequality is true by assumption, the first inequality must also be true. Hence

$$\frac{1}{2} \left(a + \frac{2}{a} \right) < a, \quad \text{if } a > 0 \text{ and } a^2 > 2.$$

Solution to Exercise D13

Rearranging the inequality (using Rule 3 and Rule 1), we obtain

$$\begin{aligned}\frac{3n}{n^2+2} < 1 &\iff 3n < n^2 + 2 \\ &\iff 0 < n^2 - 3n + 2 \\ &\iff 0 < (n-1)(n-2),\end{aligned}$$

and this final inequality certainly holds for $n > 2$. So

$$\frac{3n}{n^2+2} < 1, \quad \text{for } n > 2.$$

Solution to Exercise D14

Using Rule 3 for rearranging the inequalities, we obtain

$$2n^3 \geq (n+1)^3 \iff 2 \geq \left(\frac{n+1}{n}\right)^3.$$

The inequality on the right is certainly true for $n = 4$, since

$$\left(\frac{5}{4}\right)^3 = \frac{125}{64} < 2,$$

and, in fact, for $n \geq 4$ we have

$$\left(\frac{n+1}{n}\right)^3 = \left(1 + \frac{1}{n}\right)^3 \leq \left(1 + \frac{1}{4}\right)^3 = \left(\frac{5}{4}\right)^3,$$

so that

$$\left(\frac{n+1}{n}\right)^3 \leq 2.$$

Hence

$$2n^3 \geq (n+1)^3, \quad \text{for } n \geq 4.$$

Solution to Exercise D15

We substitute $x = 1/n$ in the Binomial Theorem for $(1+x)^n$, and notice that all the terms are positive since $n \geq 1$; this gives

$$\begin{aligned}\left(1 + \frac{1}{n}\right)^n &= 1 + n\left(\frac{1}{n}\right) + \frac{n(n-1)}{2!}\left(\frac{1}{n}\right)^2 \\ &\quad + \cdots + \left(\frac{1}{n}\right)^n \\ &\geq 1 + 1 + \frac{n-1}{2n} \\ &= 2 + \frac{1}{2} - \frac{1}{2n} = \frac{5}{2} - \frac{1}{2n}.\end{aligned}$$

Hence

$$\left(1 + \frac{1}{n}\right)^n \geq \frac{5}{2} - \frac{1}{2n}, \quad \text{for } n \geq 1.$$

Solution to Exercise D16

Let $P(n)$ be the statement $2^n \geq n^3$.

$P(10)$ is true since $2^{10} = 1024 > 10^3$.

Now let $k \geq 10$ and assume that $P(k)$ is true; that is,

$$2^k \geq k^3.$$

We wish to deduce that $P(k+1)$ is true; that is,

$$2^{k+1} \geq (k+1)^3.$$

Multiplying the inequality $2^k \geq k^3$ by 2 we get

$$2^{k+1} \geq 2k^3,$$

so it is sufficient for our purposes to prove that

$$2k^3 \geq (k+1)^3, \quad \text{for } k \geq 10.$$

This inequality is true, by Exercise D14. (It holds in fact for $k \geq 4$.) Hence

$$2^{k+1} \geq (k+1)^3, \quad \text{for } k \geq 10.$$

Thus

$$P(k) \implies P(k+1), \quad \text{for } k \geq 10.$$

Hence, by mathematical induction, $P(n)$ is true, for $n \geq 10$.

Solution to Exercise D17

Applying Bernoulli's Inequality with $x = -3/(4n)$, we obtain

$$\left(1 - \frac{3}{4n}\right)^n \geq 1 + n\left(-\frac{3}{4n}\right) = \frac{1}{4}.$$

Hence, by Rule 5 with $p = n$, we get

$$1 - \frac{3}{4n} \geq \frac{1}{4^{1/n}}, \quad \text{for } n \geq 1.$$

Using Rule 4, we can rewrite this inequality in the form

$$\frac{4n}{4n-3} \leq 4^{1/n}, \quad \text{for } n \geq 1,$$

so that

$$\frac{4n-3+3}{4n-3} = 1 + \frac{3}{4n-3} \leq 4^{1/n}, \quad \text{for } n \geq 1,$$

as required.

Next, applying Bernoulli's Inequality with $x = 3/n$, we obtain

$$\left(1 + \frac{3}{n}\right)^n \geq 1 + n\left(\frac{3}{n}\right) = 4.$$

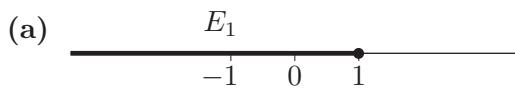
Hence, using Rule 5 with $p = n$, we get

$$1 + \frac{3}{n} \geq 4^{1/n}, \quad \text{for } n \geq 1,$$

as required.

Putting these two results together, we get the required inequality.

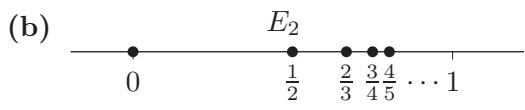
Solution to Exercise D18



The set E_1 is bounded above. For example, $M = 1$ is an upper bound of E_1 , since

$$x \leq 1, \quad \text{for all } x \in E_1.$$

Also, $\max E_1 = 1$, since $1 \in E_1$.



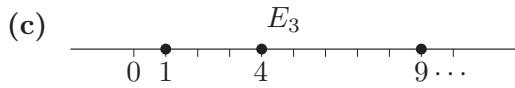
The set E_2 is bounded above. For example, $M = 1$ is an upper bound of E_2 , since

$$1 - \frac{1}{n} \leq 1, \quad \text{for } n = 1, 2, \dots$$

However, E_2 has no maximum element. For each $x \in E_2$ we have $x = 1 - 1/n$, for some $n \in \mathbb{N}$, so there is another element of E_2 for example $1 - 1/(n+1)$, such that

$$1 - \frac{1}{n} < 1 - \frac{1}{n+1} \quad \left(\text{since } \frac{1}{n+1} < \frac{1}{n}\right).$$

Hence x is not a maximum element of E_2 .



The set E_3 is not bounded above. For each real number M , there is a positive integer n such that $n^2 > M$ (for instance, take $n > M$, which implies that $n^2 \geq n > M$). Hence M is not an upper bound of E_3 .

It follows that E_3 cannot have a maximum element.

Solution to Exercise D19

(a) The set $E_1 = (-\infty, 1]$ is not bounded below. For each real number m , there is a negative real number x such that $x < m$. Since $x \in E_1$, the number m is not a lower bound of E_1 .

It follows that E_1 cannot have a minimum element.

(b) The set E_2 is bounded below by 0, since

$$1 - \frac{1}{n} \geq 0, \quad \text{for } n = 1, 2, \dots$$

Also, $0 \in E_2$, so $\min E_2 = 0$.

(c) The set E_3 is bounded below by 1, since

$$n^2 \geq 1, \quad \text{for } n = 1, 2, \dots$$

Also, $1 \in E_3$, so $\min E_3 = 1$.

Solution to Exercise D20

(a) The set $E_1 = (-\infty, 1]$ has maximum element 1, so

$$\sup E_1 = \max E_1 = 1.$$

(b) We know from Exercise D18 that 1 is an upper bound of E_2 , since

$$1 - \frac{1}{n} \leq 1, \quad \text{for } n = 1, 2, \dots$$

To show that $M = 1$ is the least upper bound of E_2 , we have to prove that, if $M' < 1$, then there is an element $1 - 1/n$ of E_2 such that

$$1 - \frac{1}{n} > M'.$$

Suppose that $0 < M' < 1$. We have

$$\begin{aligned} 1 - \frac{1}{n} &> M' \\ \iff 1 - M' &> \frac{1}{n} \\ \iff 1/(1 - M') &< n \quad (\text{since } 1 - M' > 0). \end{aligned}$$

By the Archimedean Property, we can take a positive integer n so large that $n > 1/(1 - M')$. Hence, for this value of n , we have

$$1 - \frac{1}{n} > M',$$

so M' is not an upper bound of E_2 .

It follows that 1 is the least upper bound of E_2 .

(c) The set $E_3 = \{n^2 : n = 1, 2, \dots\}$ is not bounded above, so it cannot have a least upper bound.

Solution to Exercise D21

(a) We know that 1 is a lower bound of the set $E_1 = (1, 5]$, since

$$1 \leq x, \quad \text{for all } x \in E_1.$$

To show that $m = 1$ is the greatest lower bound of E_1 , we prove that if $m' > 1$, then there is an element x in E_1 which is less than m' .

Suppose that $m' > 1$. By the Density Property, there is a real number x such that

$$1 < x < m' \quad \text{and} \quad x \leq 5,$$

so $x \in E_1$ and $x < m'$. Thus m' is not a lower bound of E_1 .

Hence 1 is the greatest lower bound of E_1 .

(b) We know that 0 is a lower bound of the set $E_2 = \{1/n^2 : n = 1, 2, \dots\}$, since

$$0 < 1/n^2, \quad \text{for } n = 1, 2, \dots$$

To show that $m = 0$ is the greatest lower bound of E_2 , we prove that if $m' > 0$, then there is an element $1/n^2$ in E_2 such that $1/n^2 < m'$.

If $m' > 0$, then we have

$$\begin{aligned} \frac{1}{n^2} < m' &\iff n^2 > \frac{1}{m'} \\ &\iff n > \sqrt{1/m'}. \end{aligned}$$

We can take a positive integer n so large that $n > \sqrt{1/m'}$. Hence the first inequality holds for this value of n , and so m' is not a lower bound of E_2 .

It follows that 0 is the greatest lower bound of E_2 .

Unit D2

Sequences

Introduction

This unit deals with sequences of real numbers, such as

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \dots,$$

$$0, 1, 0, 1, 0, 1, \dots,$$

$$1, 2, 4, 8, 16, 32, \dots.$$

The three dots (an ellipsis) indicate that the sequence continues indefinitely.

You will learn about various properties that a sequence may possess, the most important of which is *convergence*. Roughly speaking, a sequence is *convergent*, or *tends to a limit*, if the numbers in the sequence approach arbitrarily close to a unique real number, which is called the *limit* of the sequence. For example, you will see that the sequence

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \dots,$$

is convergent with limit 0. On the other hand, the numbers in the sequence

$$0, 1, 0, 1, 0, 1, \dots,$$

do not approach arbitrarily close to a unique real number, so this sequence is not convergent. Likewise, the sequence

$$1, 2, 4, 8, 16, 32, \dots,$$

is not convergent. A sequence which is not convergent is called *divergent*.

One of the reasons for studying sequences is that they provide a relatively simple setting in which we can begin to explore precise definitions of these ideas of convergence and limits. As you will see in the next unit, sequences are also a key tool in deciding when and how infinite sums make sense.

Intuitively, it seems plausible that some sequences are convergent, whereas others are not. However, the above description of convergence, involving the phrase ‘approach arbitrarily close to’, lacks the precision required in pure mathematics. If we wish to work in a serious way with convergent sequences, prove results about them and decide beyond doubt whether or not a given sequence is convergent, then we need a rigorous definition of this concept.

Historically, such a definition emerged only in the late nineteenth century, when mathematicians such as Bolzano, Cantor, Cauchy, Dedekind and Weierstrass placed analysis on a rigorous footing. It is not surprising, therefore, that at first sight the definition of convergence is rather subtle and it may take you a little time to grasp it fully.

1 Introducing sequences

In this section you will see how to picture the behaviour of a sequence by drawing a *sequence diagram*. You will also study *monotonic* sequences, that is, sequences which are either increasing or decreasing.

1.1 What is a sequence?

Ever since you learned to count, you have been familiar with the sequence of natural numbers

$$1, 2, 3, 4, 5, 6, \dots$$

You will also have encountered many other sequences of numbers, such as

$$2, 4, 6, 8, 10, 12, \dots,$$

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{64}, \dots$$

We begin our study of sequences with some definitions and notation.

Definitions

A **sequence** is an unending list of real numbers

$$a_1, a_2, a_3, \dots$$

The real number a_n is called the **n th term** of the sequence, and the sequence is denoted by

$$(a_n).$$

In the examples above, it is assumed that all the terms of the sequence after the first few are obtained by continuing the pattern in an obvious way. However, it is usually better to give a precise description of a typical term of the sequence, and often we can do this by giving an explicit formula for the n th term. For example, the n th term of the sequence (a_n) whose first few terms are

$$1, 3, 5, 7, 9, 11, \dots,$$

is given by the formula

$$a_n = 2n - 1, \quad n = 1, 2, \dots$$

We often refer to a sequence by writing the formula for its n th term in round brackets. In this notation, the sequence (a_n) would be written as $(2n - 1)$, where it is understood that n takes the successive values $1, 2, \dots$.

Remarks

- Although most texts on analysis use the notation (a_n) for a sequence, you may also come across the alternative notations $\{a_n\}$ or $\langle a_n \rangle$.
- Of course, round brackets are used frequently in mathematics with a variety of different meanings, so there is some risk of ambiguity in using the notation (a_n) for a sequence. This is especially true when we refer to a sequence by writing the formula for its n th term in round brackets, as with the sequence $(2n - 1)$ mentioned above. Usually, the context will make clear whether or not an expression in round brackets refers to a sequence. For example, if it appears in the middle of an equation, an expression in round brackets does *not* denote a sequence. Thus in the equation $2n^2 - n = n(2n - 1)$, the expression $(2n - 1)$ does not refer to a sequence.
- Notice that a sequence of real numbers differs from a *set* of real numbers. Changing the order of the terms in a sequence gives us a new sequence, whereas rearranging the elements of a set leaves the set unchanged. Moreover, the same number can occur many times in a sequence, but not in a set; for example, all the terms in the sequence

$$0, 1, 0, 1, 0, 1, \dots,$$

belong to the set $\{0, 1\}$. One advantage of the round bracket notation for sequences is that it avoids any confusion with sets that might arise from using braces (curly brackets).

- For all the sequences you will meet in this unit, there will be an explicit formula for the n th term. However, this is not essential – for example, the sequence of digits in the decimal expansion of π is a well-defined sequence, but there is no formula for its n th term.

Exercise D22

Calculate the first five terms of each of the following sequences (a_n) .

(Give your answer to part (e) to two decimal places.)

- $a_n = 3n + 1$, $n = 1, 2, \dots$
- $a_n = 3^{-n}$, $n = 1, 2, \dots$
- $a_n = (-1)^n n$, $n = 1, 2, \dots$
- $a_n = n!$, $n = 1, 2, \dots$
- $a_n = \left(1 + \frac{1}{n}\right)^n$, $n = 1, 2, \dots$

Sequences often begin with a term corresponding to $n = 1$. However, sometimes it is necessary (or convenient) to begin a sequence with some other value of n . For example, the sequence (a_n) defined by

$$a_n = 1/(n! - n)$$

cannot begin with $n = 1$ or 2. We indicate this by writing, for example, $(a_n)_3^\infty$ to represent the sequence

$$a_3, a_4, a_5, \dots$$

or writing

$$a_n = 1/(n! - n), \quad n = 3, 4, \dots$$

If a sequence is written as (a_n) with no subscripts or superscripts, then we assume that this denotes the sequence $(a_n)_1^\infty$.

Sequence diagrams

One helpful way to think of a sequence is as a *function* f with domain \mathbb{N} and codomain \mathbb{R} that maps each number n in the domain to the n th term of the sequence:

$$\begin{aligned} f : \mathbb{N} &\longrightarrow \mathbb{R} \\ n &\longmapsto a_n \end{aligned}$$

Using this idea, we can picture how a given sequence (a_n) behaves by drawing its **sequence diagram**, that is, the graph of the function from \mathbb{N} to \mathbb{R} that represents the sequence. To do this, we mark suitable values of n on the horizontal axis and, for each value of n , we plot the point (n, a_n) .

Often it is necessary to use different scales on the axes for clarity. Figure 1 shows the sequence diagrams for three different sequences.

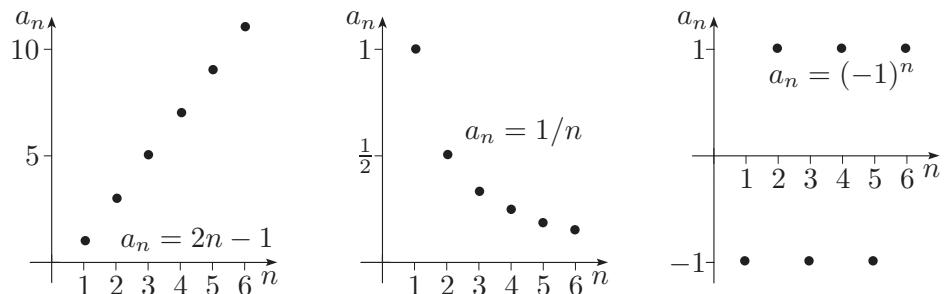


Figure 1 Three sequence diagrams

Exercise D23

Draw a sequence diagram, showing the first five points, for each of the following sequences (a_n) .

(In part (c) you can use your solution to Exercise D22(e).)

- (a) $a_n = n^2$, $n = 1, 2, \dots$
- (b) $a_n = 3$, $n = 1, 2, \dots$
- (c) $a_n = \left(1 + \frac{1}{n}\right)^n$, $n = 1, 2, \dots$
- (d) $a_n = \frac{(-1)^n}{n}$, $n = 1, 2, \dots$

1.2 Monotonic sequences

Many sequences have the property that, as n increases, their terms are either *increasing* or *decreasing*. For example, the sequence $(2n - 1)$ has terms $1, 3, 5, 7, \dots$, which are increasing, whereas the sequence $(1/n)$ has terms $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$, which are decreasing. The sequence $((-1)^n)$ is neither increasing nor decreasing. All this can be seen clearly on the sequence diagrams in Figure 1.

We now give precise meanings to these words *increasing* and *decreasing*, and introduce the word *monotonic*. These terms are illustrated in Figure 2.

Definitions

A sequence (a_n) is said to be

- **constant** if

$$a_{n+1} = a_n, \quad \text{for } n = 1, 2, \dots;$$

- **increasing** if

$$a_{n+1} \geq a_n, \quad \text{for } n = 1, 2, \dots,$$

and **strictly increasing** if

$$a_{n+1} > a_n, \quad \text{for } n = 1, 2, \dots;$$

- **decreasing** if

$$a_{n+1} \leq a_n, \quad \text{for } n = 1, 2, \dots,$$

and **strictly decreasing** if

$$a_{n+1} < a_n, \quad \text{for } n = 1, 2, \dots;$$

- **monotonic** if (a_n) is either increasing or decreasing.

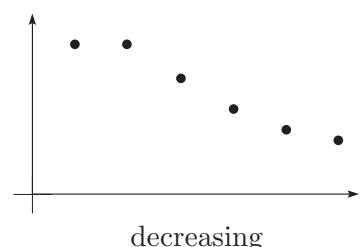
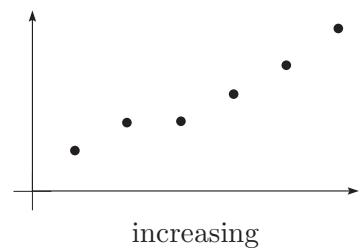
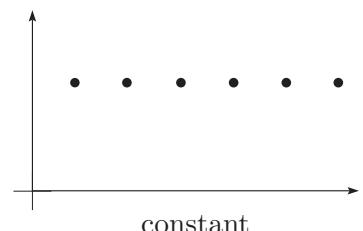


Figure 2 Monotonic sequences

Notice that for a sequence (a_n) to be *increasing*, it is essential that $a_{n+1} \geq a_n$ for all $n \geq 1$. However, we do not require strict inequalities because we wish to describe a sequence such as

$$1, 1, 2, 2, 3, 3, 4, 4, \dots$$

as increasing. If strict inequalities do hold, and we want to emphasise this, then we can use the term *strictly increasing* to describe the sequence, provided that $a_{n+1} > a_n$ for all $n \geq 1$. Thus every strictly increasing sequence is increasing, but the converse is not true. Similar comments apply to the terms *decreasing* and *strictly decreasing*. One slightly bizarre consequence of the definitions is that constant sequences are both increasing and decreasing (though not, of course, strictly increasing or strictly decreasing).

A sequence is *monotonic* if it is either increasing or decreasing (we do not require it to be strictly increasing or strictly decreasing, although it may be). To determine whether a given sequence is monotonic, it is not sufficient to draw a diagram: it is necessary to give a proof. There are various ways to do this. For example, $(1/n)$ is a strictly decreasing sequence and is therefore monotonic, because

$$\frac{1}{n+1} < \frac{1}{n}, \quad \text{for } n = 1, 2, \dots,$$

since $n+1 > n > 0$, for $n = 1, 2, \dots$. Here we have used Rule 4 for rearranging inequalities that you met in Unit D1 *Numbers*. We often use these rules when dealing with sequences, so for convenience they are restated in the box below.

Rules for rearranging inequalities

Let a, b, c and p be real numbers.

Rule 1 $a < b \iff b - a > 0$.

Rule 2 $a < b \iff a + c < b + c$.

Rule 3 If $c > 0$, then $a < b \iff ac < bc$;
if $c < 0$, then $a < b \iff ac > bc$.

Rule 4 If $a, b > 0$, then

$$a < b \iff \frac{1}{a} > \frac{1}{b}.$$

Rule 5 If $a, b \geq 0$ and $p > 0$, then

$$a < b \iff a^p < b^p.$$

Rule 6 $|a| < b \iff -b < a < b$.

All the above rules also hold if strict inequalities are replaced by weak inequalities.

The next worked exercise illustrates some of the other approaches that can be used to prove whether or not a sequence is monotonic.

Worked Exercise D19

Determine which of the following sequences (a_n) are monotonic.

- (a) $a_n = \frac{1}{n}$, $n = 1, 2, \dots$
- (b) $a_n = (n-1)(n-2)$, $n = 1, 2, \dots$
- (c) $a_n = (-1)^n$, $n = 1, 2, \dots$

Solution

- (a) You have just seen one proof that this sequence is monotonic. Here is another.

We have

$$a_n = 1/n \quad \text{and} \quad a_{n+1} = 1/(n+1),$$

so

$$\frac{a_{n+1}}{a_n} = \frac{1/(n+1)}{1/n} = \frac{n}{n+1} < 1, \quad \text{for } n = 1, 2, \dots$$

It follows that

$$a_{n+1} < a_n, \quad \text{for } n = 1, 2, \dots$$

Thus (a_n) is strictly decreasing, so (a_n) is monotonic.

- (b) For each $n \geq 1$, we have

$$a_n = (n-1)(n-2) = n^2 - 3n + 2$$

and

$$a_{n+1} = n(n-1) = n^2 - n,$$

so

$$a_{n+1} - a_n = 2n - 2 \geq 0, \quad \text{for } n = 1, 2, \dots$$

It follows that

$$a_{n+1} \geq a_n, \quad \text{for } n = 1, 2, \dots$$

Thus (a_n) is increasing, so (a_n) is monotonic.

- Notice that this sequence is not strictly increasing, because $a_1 = a_2 = 0$.

- (c) To prove that a sequence is not monotonic, use consecutive terms to show that the sequence is neither increasing nor decreasing.

Consider the first three terms of the sequence: $a_1 = -1$, $a_2 = 1$ and $a_3 = -1$.

We have $a_3 < a_2$, which means that (a_n) is not increasing. Also $a_2 > a_1$, which means that (a_n) is not decreasing.

Thus (a_n) is neither increasing nor decreasing and so is not monotonic.

Worked Exercise D19 illustrates the use of the following two strategies.

Strategy D3

To show that a given sequence (a_n) is monotonic, consider the difference $a_{n+1} - a_n$.

- If $a_{n+1} - a_n \geq 0$, for $n = 1, 2, \dots$, then (a_n) is increasing.
- If $a_{n+1} - a_n \leq 0$, for $n = 1, 2, \dots$, then (a_n) is decreasing.

If $a_n > 0$ for all n , then it is often more convenient to use the following strategy.

Strategy D4

To show that a given sequence (a_n) of *positive* terms is monotonic, consider the quotient $\frac{a_{n+1}}{a_n}$.

- If $\frac{a_{n+1}}{a_n} \geq 1$, for $n = 1, 2, \dots$, then (a_n) is increasing.
- If $\frac{a_{n+1}}{a_n} \leq 1$, for $n = 1, 2, \dots$, then (a_n) is decreasing.

For a positive sequence, you can use either strategy; which is best depends on whether you think it is easier to simplify the difference $a_{n+1} - a_n$ or the quotient a_{n+1}/a_n .

Exercise D24

Show that the following sequences (a_n) are monotonic.

(In part (a) remember that, by convention, $0! = 1$.)

- $a_n = (n - 1)!$, $n = 1, 2, \dots$
- $a_n = 2^{-n}$, $n = 1, 2, \dots$
- $a_n = n + \frac{1}{n}$, $n = 1, 2, \dots$

Often it is possible to *guess* whether or not a sequence defined by a formula is monotonic by calculating the first few terms. Consider, for example, the sequence (a_n) given by

$$a_n = \left(1 + \frac{1}{n}\right)^n, \quad n = 1, 2, \dots$$

In Exercise D22(e) you found that the first five terms of this sequence are approximately

$$2, 2.25, 2.37, 2.44, 2.49.$$

These terms suggest that the sequence (a_n) is increasing and, in fact, it is, as you will see in Section 5.

However, the first few terms of a sequence are not always a reliable guide to the sequence's behaviour. Consider, for example, the sequence

$$a_n = \frac{10^n}{n!}, \quad n = 1, 2, \dots$$

The first five terms of this sequence are approximately

$$10, 50, 167, 417, 833.$$

These terms suggest that (a_n) is increasing. However, calculation of more terms shows that this is not so, as you can see in Figure 3.

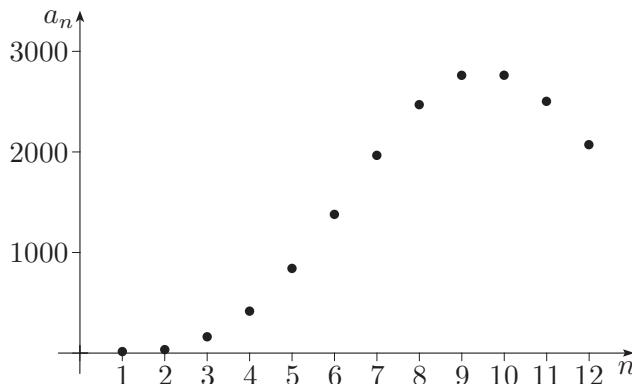


Figure 3 The sequence diagram for $a_n = 10^n/n!$

If we use Strategy D4, we find that

$$\frac{a_{n+1}}{a_n} = \frac{10^{n+1}/(n+1)!}{10^n/n!} = \frac{10^{n+1}n!}{10^n(n+1)!} = \frac{10}{n+1}.$$

Now

$$\begin{aligned} \frac{10}{n+1} \leq 1 &\iff n+1 \geq 10 \\ &\iff n \geq 9. \end{aligned}$$

So

$$\frac{a_{n+1}}{a_n} \leq 1, \quad \text{for } n \geq 9.$$

Hence

$$a_{n+1} \leq a_n, \quad \text{for } n \geq 9,$$

so (a_n) is eventually decreasing (in fact, $a_9 = a_{10}$ and (a_n) is strictly decreasing for $n \geq 10$).

This type of situation arises quite often, so it is helpful to give a formal definition.

Definition

If a sequence (a_n) has a certain property provided we ignore a finite number of terms, we say that the sequence **eventually** has this property.

We have just seen that the sequence $(10^n/n!)$ is eventually decreasing. As another example of this usage, consider the sequence (a_n) defined by the formula

$$a_n = n^2, \quad n = 1, 2, \dots$$

Then we can say that the terms of this sequence are eventually greater than 100, because

$$n^2 > 100, \quad \text{for } n > 10.$$

Sometimes you may need to show that a sequence does *not* eventually have a certain property. To do this, you need to find infinitely many terms of the sequence which fail to have the property. For example, the terms of the sequence (a_n) defined by the formula

$$a_n = 3n, \quad n = 1, 2, \dots$$

are not eventually even, because $3n$ is an odd number whenever n is odd.

Exercise D25

Classify each of the following statements as true or false and justify your answers (if a statement is true, then prove it; if a statement is false, then explain why).

- (a) The terms of the sequence (a_n) defined by $a_n = 2^n$ are eventually greater than 1000.
- (b) The terms of the sequence (a_n) defined by $a_n = (-1)^n$ are eventually positive.
- (c) The terms of the sequence (a_n) defined by $a_n = 1/n$ are eventually less than 0.025.
- (d) The sequence (a_n) defined by $a_n = n^4/4^n$ is eventually decreasing.

2 Null sequences

In this section you will meet the definition of a *null sequence*, that is, a sequence which converges to 0. You will then explore the properties of null sequences and see how to prove these. You will also meet some basic null sequences and learn how to identify new null sequences.

2.1 What is a null sequence?

Let us begin by looking at the size of the terms of the sequence (a_n) defined by

$$a_n = \frac{1}{n}, \quad n = 1, 2, \dots$$

Worked Exercise D20

For each of the following statements about the terms of the above sequence (a_n) , find an integer N that makes the statement true.

(a) $\frac{1}{n} < \frac{1}{100}$, for all $n > N$

(b) $\frac{1}{n} < \frac{3}{1000}$, for all $n > N$

Solution

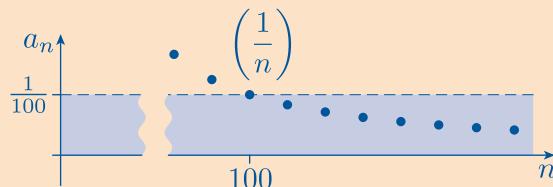
- (a) Since the quantities involved are all positive, we can use Rule 4 to rearrange the inequality.

We have that

$$\frac{1}{n} < \frac{1}{100} \iff n > 100.$$

Hence we may take $N = 100$.

This is illustrated in the sequence diagram below.



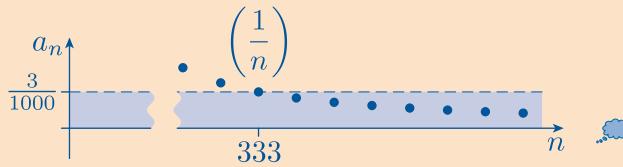
Of course, any integer greater than 100 is also a valid value for N , but any integer less than 100 is not.

- (b) We have that

$$\frac{1}{n} < \frac{3}{1000} \iff n > 333.333\dots$$

Hence we may take $N = 333$.

 The value $N = 333$ is valid since the smallest value of n that this allows is $n = 334$, and of course $334 > 333.333 \dots$. This is illustrated in the sequence diagram below.



The next exercise is similar to Worked Exercise D20, except that this time you are asked to look at a sequence with both positive and negative terms. The statements about the size of the terms therefore involve the *modulus* of the terms.

Exercise D26

For each of the following statements about the sequence (a_n) defined by the formula

$$a_n = \frac{(-1)^n}{n^2}, \quad n = 1, 2, \dots,$$

find an integer N that makes the statement true.

(a) $\left| \frac{(-1)^n}{n^2} \right| < \frac{1}{100}$, for all $n > N$

(b) $\left| \frac{(-1)^n}{n^2} \right| < \frac{3}{1000}$, for all $n > N$

The solutions of Worked Exercise D20 and Exercise D26 suggest that, if (a_n) is a sequence defined by

$$a_n = \frac{1}{n}, \quad n = 1, 2, \dots,$$

or by

$$a_n = \frac{(-1)^n}{n^2}, \quad n = 1, 2, \dots,$$

then as n becomes larger and larger, the terms of the sequence get closer and closer to 0.

We need a formal way of describing precisely what we mean by this. In order to do this we introduce the Greek letter ε , pronounced ‘epsilon’, which we use to denote a positive number that may be as small as we please in any given particular instance.

The symbol ε was first used within proofs in analysis by Augustin-Louis Cauchy (1789–1857) in his *Cours d'Analyse* of 1821. Cauchy chose ε , which he also used in some of his work on probability, because it corresponds to the initial letter of erreur (error), a fact which seems rather amusing today given that ε is now the characteristic symbol of precision and rigour in analysis.

We see that, for each of the two sequences we are considering, the terms of the sequence eventually lie inside a horizontal strip in the sequence diagram from $-\varepsilon$ up to ε , and this is the case no matter how small ε is taken to be.

For the sequence defined by

$$a_n = \frac{1}{n}, \quad n = 1, 2, \dots,$$

the sequence diagram is given again in Figure 4. In this case the terms of the sequence are all positive, so we need only look at a horizontal strip from 0 up to ε . In Worked Exercise D20 we considered the particular values $\varepsilon = 1/100$ and $\varepsilon = 3/1000$, but now we let ε represent *any* positive number, however small.

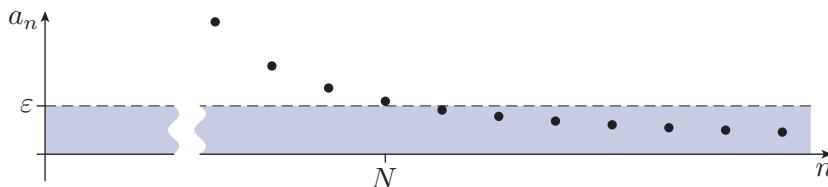


Figure 4 The sequence $a_n = 1/n$

This diagram suggests that, for each positive number ε , there is an integer N such that

$$|a_n| = a_n = \frac{1}{n} < \varepsilon, \quad \text{for all } n > N.$$

This means that every term to the right of N in the diagram lies within the horizontal strip. In fact, this will be true if we take N to be any integer satisfying $N \geq \frac{1}{\varepsilon}$.

The sequence diagram in the case that

$$a_n = \frac{(-1)^n}{n^2}, \quad n = 1, 2, \dots$$

is given in Figure 5.

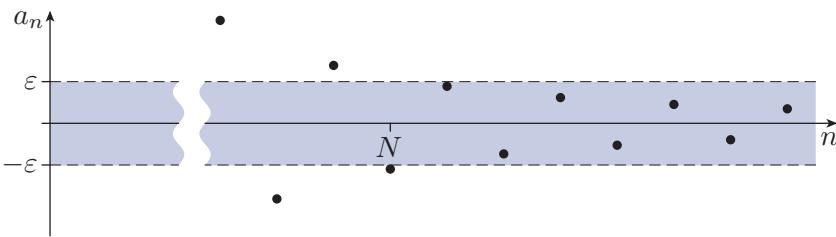


Figure 5 The sequence $a_n = \frac{(-1)^n}{n^2}$

This diagram suggests that, for each positive number ε , there is an integer N such that

$$|a_n| = \left| \frac{(-1)^n}{n^2} \right| = \frac{1}{n^2} < \varepsilon, \quad \text{for all } n > N.$$

In fact, this will be true if we take N to be any integer satisfying $N \geq \sqrt{1/\varepsilon}$.

In both cases, the smaller we choose ε , the further to the right in the sequence diagram we have to go before we can be sure that all the terms of the sequence from that point onwards lie inside the strip. That is, the smaller we choose ε the larger we have to choose N if we wish to have

$$|a_n| < \varepsilon, \quad \text{for all } n > N.$$

We now give a definition of a *null sequence* which formalises the notion of a sequence ‘getting closer and closer to 0’. It follows from the discussion above that both of the sequences we have been considering are null sequences according to this definition. The concept of a null sequence is illustrated in Figure 6.

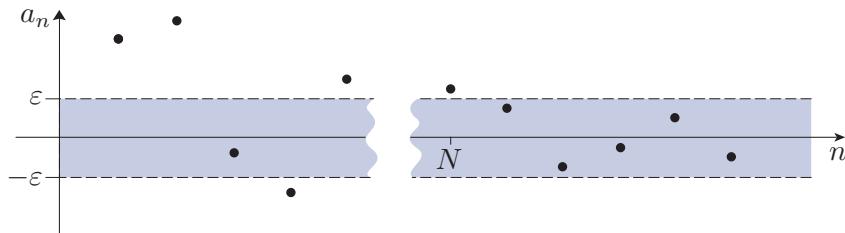


Figure 6 A null sequence

Definitions

The sequence (a_n) is **null** if

for each positive number ε , there is an integer N such that

$$|a_n| < \varepsilon, \quad \text{for all } n > N.$$

We also say that the sequence (a_n) is **convergent with limit 0**, or that (a_n) **converges** to 0.

Remarks

- We write ‘for all $n > N$ ’ to emphasise that the inequality $|a_n| < \varepsilon$ holds for every integer $n > N$. Note that we can rewrite the last line of the definition as the implication

if $n > N$, then $|a_n| < \varepsilon$.

We sometimes refer to this statement as the ε - N statement.

- The sequence (a_n) is null if and only if the sequence $(|a_n|)$ is null. This is because the statement in the definition is identical for the sequences (a_n) and $(|a_n|)$.
- The null sequence (a_n) remains null if we add, delete or alter a finite number of terms to produce a new sequence (b_n) . Informally, we say that ‘finitely many terms do not matter’.

This is because the statement in the definition above and its corresponding version for (b_n) are identical, except that the values of N may differ by some integer.

We can interpret the task of finding a suitable integer N when using the definition as an ‘ ε - N game’ in which player A chooses a positive number ε and challenges player B to find some integer N for which the statement in the definition is true. Thus, for example, it follows from Exercise D26 that, for the sequence (a_n) defined by

$$a_n = \frac{(-1)^n}{n^2}, \quad n = 1, 2, \dots,$$

if player A chooses $\varepsilon = 1/100$ then player B can choose $N = 10$ (or any larger value). This is illustrated in Figure 7.

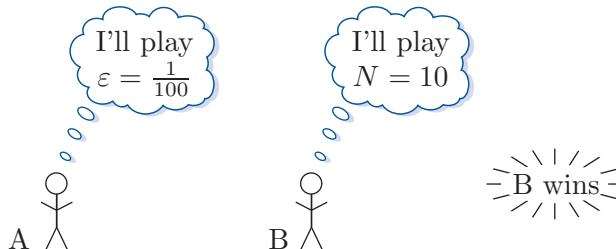


Figure 7 The ε - N game

Notice that in the ‘ ε - N game’, if (a_n) is any null sequence and both players make their choices carefully, then player B will always win.

Worked Exercise D21

Prove that (a_n) is a null sequence if

$$a_n = \frac{1}{n^3}, \quad n = 1, 2, \dots$$

Solution

We have to prove that, for each $\varepsilon > 0$, there is an integer N such that

$$\left| \frac{1}{n^3} \right| < \varepsilon, \quad \text{for all } n > N. \quad (*)$$

 In order to find a suitable value of N , we rearrange the inequality in $(*)$ until we obtain an inequality with just n on one side. We use Rules 4 and 5 for rearranging inequalities. 

We have that

$$\begin{aligned} \left| \frac{1}{n^3} \right| < \varepsilon &\iff \frac{1}{n^3} < \varepsilon \\ &\iff n^3 > \frac{1}{\varepsilon} \\ &\iff n > \frac{1}{\sqrt[3]{\varepsilon}}. \end{aligned}$$

So, statement $(*)$ holds if $N \geq \frac{1}{\sqrt[3]{\varepsilon}}$. Hence (a_n) is null.

Sometimes we might want to prove that a sequence (a_n) is *not* null. To do this, we have to show that (a_n) does not satisfy the definition of a null sequence. In other words, we must show that the following statement is true:

there is some value of $\varepsilon > 0$ for which there is no integer N such that

$$|a_n| < \varepsilon, \quad \text{for all } n > N.$$

This is illustrated in the next worked exercise. Notice that in the ' ε - N game', if (a_n) is not a null sequence and both players make their choices carefully, then player A will always win.

Worked Exercise D22

Prove that the following sequence (a_n) is not null:

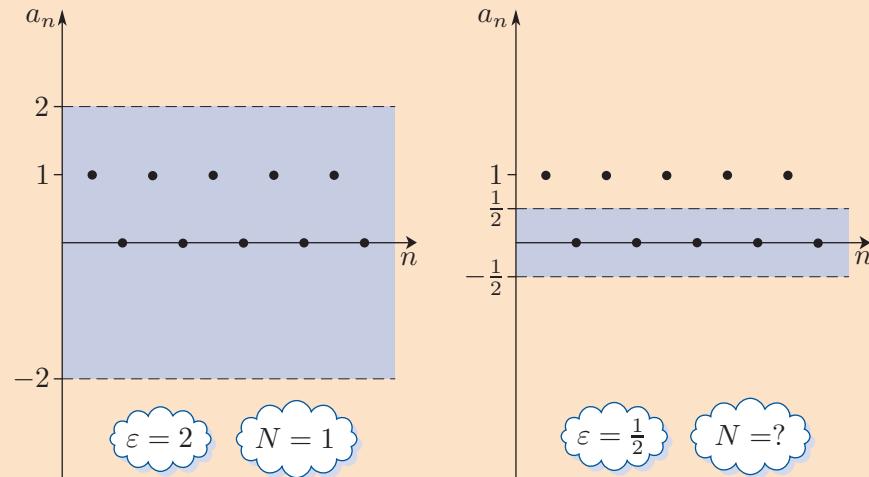
$$a_n = \begin{cases} 1, & \text{if } n \text{ is odd,} \\ 0, & \text{if } n \text{ is even.} \end{cases}$$

Solution

We have to find a positive number ε for which there is no integer N such that

$$|a_n| < \varepsilon, \quad \text{for all } n > N.$$

In terms of the sequence diagram, this means that the sequence (a_n) does not eventually lie in the horizontal strip from $-\varepsilon$ to ε .



We can see that 2 would not be a suitable value of ε but $\frac{1}{2}$ would be a suitable value. In our ' ε - N game', if player A plays $\varepsilon = \frac{1}{2}$, then there is no integer N that player B can play to win. ☺

For all odd values of n , we have $|a_n| = 1$ and so, if $\varepsilon = \frac{1}{2}$, there is no integer N such that

$$|a_n| < \varepsilon, \quad \text{for all } n > N.$$

Hence (a_n) is not a null sequence.

Notice that any positive value of ε less than 1 will serve our purpose here: there is nothing special about the number $\frac{1}{2}$. ☺

Worked Exercises D21 and D22 illustrate the following strategy.

Strategy D5

- To show that the sequence (a_n) is null, rearrange the inequality $|a_n| < \varepsilon$ to find an integer N (generally depending on ε) such that $|a_n| < \varepsilon$, for all $n > N$.
- To show that the sequence (a_n) is not null, find *one* value of $\varepsilon > 0$ for which there is *no* integer N such that $|a_n| < \varepsilon$ for all $n > N$.

Exercise D27

Use Strategy D5 to determine which of the following sequences (a_n) are null.

$$(a) \quad a_n = \frac{1}{2n-1}, \quad n = 1, 2, \dots$$

$$(b) \quad a_n = \frac{(-1)^n}{10}, \quad n = 1, 2, \dots$$

$$(c) \quad a_n = \frac{(-1)^n}{n^4 + 1}, \quad n = 1, 2, \dots$$

Hint: You will need to consider the case where $\varepsilon \geq 1$ and the case where $0 < \varepsilon < 1$ separately.

2.2 Properties of null sequences

We now look at a number of properties of null sequences. These allow us to identify new null sequences without having to work with the definition.

This subsection contains several proofs; reading them should improve your understanding of the material. However, if you are short of time, you should skim through these proofs now and return to them when time permits.

Theorem D4 Power Rule for null sequences

If (a_n) is null, where $a_n \geq 0$, for $n = 1, 2, \dots$, and p is a positive real number, then (a_n^p) is null.

Proof We want to prove that the sequence (a_n^p) is null; that is:

for each positive number ε , there is an integer N such that

$$a_n^p < \varepsilon, \quad \text{for all } n > N. \quad (1)$$

Here we use the fact that $|a_n^p| = a_n^p$, since $a_n \geq 0$.

We start by letting ε be a positive number. Since (a_n) is null and $\varepsilon^{1/p}$ is positive, there is an integer N such that

$$a_n < \varepsilon^{1/p}, \quad \text{for all } n > N. \quad (2)$$

Taking the p th power of both sides of the inequality in statement (2), we see that statement (1) holds with the same value of N . ■

Note how we used $\varepsilon^{1/p}$ in statement (2) in order to obtain ε in statement (1). We often prove the ε - N statement for some new null sequence by applying the definition to a known null sequence (or sequences), using a positive number related in a suitable way to ε .

Earlier we saw that the sequence (a_n) defined by

$$a_n = \frac{1}{n}, \quad n = 1, 2, \dots$$

is null. By applying the Power Rule with $p = 3$, we can now deduce that the sequence (b_n) defined by

$$b_n = \frac{1}{n^3}, \quad n = 1, 2, \dots$$

is also null.

We will use the next set of rules a great deal.

Theorem D5 Combination Rules for null sequences

If (a_n) and (b_n) are null, then:

Sum Rule $(a_n + b_n)$ is null

Multiple Rule (λa_n) is null, for any real number λ

Product Rule $(a_n b_n)$ is null.

Proof We first prove the Sum Rule. We want to prove that the sequence $(a_n + b_n)$ is null; that is:

for each positive number ϵ , there is an integer N such that

$$|a_n + b_n| < \epsilon, \quad \text{for all } n > N. \quad (3)$$

Let ϵ be a positive number. Since (a_n) and (b_n) are null, there are integers N_1 and N_2 such that

$$|a_n| < \frac{1}{2}\epsilon, \quad \text{for all } n > N_1, \quad \text{and} \quad |b_n| < \frac{1}{2}\epsilon, \quad \text{for all } n > N_2.$$

 We use $\frac{1}{2}\epsilon$ here in order to obtain ϵ in statement (3). 

If $N = \max\{N_1, N_2\}$, then both the above inequalities hold for all $n > N$. Therefore, by the Triangle Inequality (which you met in Unit D1),

$$|a_n + b_n| \leq |a_n| + |b_n| < \frac{1}{2}\epsilon + \frac{1}{2}\epsilon = \epsilon, \quad \text{for all } n > N.$$

Thus statement (3) holds with this value of N .

Next, we prove the Multiple Rule. We want to prove that the sequence (λa_n) is null; that is:

for each positive number ϵ , there is an integer N such that

$$|\lambda a_n| < \epsilon, \quad \text{for all } n > N. \quad (4)$$

If $\lambda = 0$, this statement is obvious, so we can assume that $\lambda \neq 0$.

Let ϵ be a positive number. Since (a_n) is null, there is an integer N such that

$$|a_n| < \epsilon/|\lambda|, \quad \text{for all } n > N.$$

 We use $\epsilon/|\lambda|$ here in order to obtain ϵ in statement (4). 

Multiplying both sides of this inequality by the positive number $|\lambda|$ gives

$$|\lambda a_n| < \epsilon, \quad \text{for all } n > N.$$

Thus statement (4) holds with this value of N .

Finally, we prove the Product Rule. We want to prove that the sequence $(a_n b_n)$ is null; that is:

for each positive number ε , there is an integer N such that

$$|a_n b_n| < \varepsilon, \quad \text{for all } n > N. \quad (5)$$

Let ε be a positive number. Since (a_n) and (b_n) are null, there are integers N_1 and N_2 such that

$$|a_n| < \sqrt{\varepsilon}, \quad \text{for all } n > N_1, \quad \text{and} \quad |b_n| < \sqrt{\varepsilon}, \quad \text{for all } n > N_2.$$

We use $\sqrt{\varepsilon}$ here in order to obtain ε in statement (5).

If $N = \max\{N_1, N_2\}$, then both the above inequalities hold for all $n > N$, so if we multiply them we obtain

$$|a_n b_n| = |a_n||b_n| < \sqrt{\varepsilon}\sqrt{\varepsilon} = \varepsilon, \quad \text{for all } n > N.$$

Thus statement (5) holds with this value of N . ■

Worked Exercise D23

Use the Power and Combination Rules, and any sequences that you have already shown to be null, to show that the following sequences (a_n) are null.

$$(a) \quad a_n = \frac{1}{n} + \frac{1}{n^3}, \quad n = 1, 2, \dots$$

$$(b) \quad a_n = \frac{6}{\sqrt[5]{n}} + \frac{5}{(2n-1)^7}, \quad n = 1, 2, \dots$$

$$(c) \quad a_n = \frac{1}{n(2n-1)}, \quad n = 1, 2, \dots$$

Solution

(a) We know that the sequences $\left(\frac{1}{n}\right)$ and $\left(\frac{1}{n^3}\right)$ are null, so (a_n) is null, by the Sum Rule.

(b) The sequences $\left(\frac{1}{n}\right)$ and $\left(\frac{1}{2n-1}\right)$ are null, so the sequences $\left(\frac{6}{\sqrt[5]{n}}\right)$ and $\left(\frac{5}{(2n-1)^7}\right)$ are null, by the Power Rule and the Multiple Rule.

Hence (a_n) is null, by the Sum Rule.

(c) We know that $\left(\frac{1}{n}\right)$ and $\left(\frac{1}{2n-1}\right)$ are null, so (a_n) is null, by the Product Rule.

Exercise D28

Use the Power and Combination Rules, and any sequences that you have already shown to be null, to show that the following sequences (a_n) are null.

$$(a) \quad a_n = \frac{1}{(2n-1)^3}, \quad n = 1, 2, \dots$$

$$(b) \quad a_n = \frac{7\pi}{n^3}, \quad n = 1, 2, \dots$$

$$(c) \quad a_n = \frac{1}{3n^4(2n-1)^{1/3}}, \quad n = 1, 2, \dots$$

Our next rule, the Squeeze Rule, also enables us to ‘get new null sequences from old’ – but in a slightly different way. To illustrate this rule, we look first at the sequence diagrams of two sequences (a_n) and (b_n) defined by

$$a_n = \frac{1}{1 + \sqrt{n}}, \quad n = 1, 2, \dots$$

and

$$b_n = \frac{1}{\sqrt{n}}, \quad n = 1, 2, \dots$$

We know that (b_n) is a null sequence by the Power Rule, but what about (a_n) ?

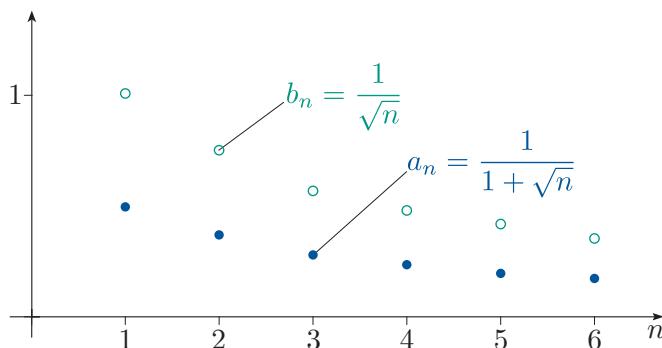


Figure 8 The sequence diagrams for $a_n = \frac{1}{1 + \sqrt{n}}$ and $b_n = \frac{1}{\sqrt{n}}$

Figure 8 shows that the points corresponding to the sequence (a_n) are ‘squeezed’ in between the horizontal axis and the points corresponding to the null sequence (b_n) , since

$$0 < \frac{1}{1 + \sqrt{n}} < \frac{1}{\sqrt{n}}, \quad \text{for } n = 1, 2, \dots$$

We express this in words by saying that the sequence (a_n) is **dominated** by the sequence (b_n) . Hence, if from some point onwards all the points corresponding to (b_n) lie in a narrow horizontal strip in the sequence diagram from $-\varepsilon$ up to ε , then (from the same point onwards) all the points corresponding to (a_n) will also lie in the same strip. So, since ε may be any positive number, it certainly looks from this sequence diagram argument that (a_n) must be a null sequence too.

Theorem D6 Squeeze Rule for null sequences

If (b_n) is a null sequence of non-negative terms, and

$$|a_n| \leq b_n, \quad \text{for } n = 1, 2, \dots,$$

then (a_n) is null.

Proof We want to prove that (a_n) is null; that is:

for each positive number ε , there is an integer N such that

$$|a_n| < \varepsilon, \quad \text{for all } n > N. \quad (6)$$

Let ε be a positive number. Then since (b_n) is null and its terms are non-negative, there is some integer N such that

$$|b_n| = b_n < \varepsilon, \quad \text{for all } n > N.$$

We also know that $|a_n| \leq b_n$, for $n = 1, 2, \dots$, so it follows that

$$|a_n| < \varepsilon, \quad \text{for all } n > N.$$

Thus statement (6) holds with this value of N . ■

To show that a sequence is null using the Squeeze Rule, we use the following strategy.

Strategy D6

To use the Squeeze Rule to show that a sequence (a_n) is null, do the following.

1. Guess a dominating null sequence (b_n) with non-negative terms.
2. Check that $|a_n| \leq b_n$, for $n = 1, 2, \dots$.

The following worked exercise illustrates the use of the strategy.

Worked Exercise D24

Use the Squeeze Rule to show that the following sequences (a_n) are null.

$$(a) \quad a_n = \frac{(-1)^n}{n^3 + 1}, \quad n = 1, 2, \dots$$

$$(b) \quad a_n = \frac{2 \cos(2n)}{n^2}, \quad n = 1, 2, \dots$$

Solution

- (a) To guess a suitable dominating sequence, we look at the formula for a_n and try to spot a related null sequence whose terms have larger magnitude and are non-negative. Here, $(1/n^3)$ seems to be a likely candidate.

We guess that (a_n) is dominated by (b_n) , where

$$b_n = \frac{1}{n^3}, \quad n = 1, 2, \dots$$

To check this, we have to show that

$$\left| \frac{(-1)^n}{n^3 + 1} \right| \leq \frac{1}{n^3}, \quad \text{for } n = 1, 2, \dots$$

This holds because

$$n^3 + 1 \geq n^3, \quad \text{for } n = 1, 2, \dots$$

We showed earlier that (b_n) is null, so we can deduce that (a_n) is null, by the Squeeze Rule.

- (b) In this case, we know that $-1 \leq \cos(2n) \leq 1$ by the properties of the cosine function. This suggests that $(2/n^2)$ might be a suitable dominating sequence.

We guess that (a_n) is dominated by (b_n) , where

$$b_n = \frac{2}{n^2}, \quad n = 1, 2, \dots$$

To check this, we have to show that

$$\left| \frac{2 \cos(2n)}{n^2} \right| \leq \frac{2}{n^2}, \quad \text{for } n = 1, 2, \dots$$

This holds because

$$|\cos(2n)| \leq 1, \quad \text{for } n = 1, 2, \dots$$

We showed earlier that the sequence $(1/n)$ is null, so it follows from the Power Rule and the Multiple Rule that (b_n) is null. We deduce that (a_n) is null, by the Squeeze Rule.

Exercise D29

Use the Squeeze Rule to show that the following sequences (a_n) are null.

(a) $a_n = \frac{1}{n^2 + n}, \quad n = 1, 2, \dots$

(b) $a_n = \frac{(-1)^n}{n!}, \quad n = 1, 2, \dots$

(c) $a_n = \frac{\sin(n^2)}{n^2 + 2^n}, \quad n = 1, 2, \dots$

2.3 Basic null sequences

We now show that there are various basic types of sequences that are null. By applying the rules from the previous subsection to these ‘basic null sequences’, we can deduce the existence of many different null sequences without having to use the definition.

It is important that you are familiar with these types of basic null sequences and are able to use them. Reading the proof that they are null may help you with this, but skim read it if you are short of time and return to it when time permits.

Theorem D7 Basic null sequences

The following sequences are null.

- (a) $(1/n^p)$, for $p > 0$.
- (b) (c^n) , for $|c| < 1$.
- (c) $(n^p c^n)$, for $p > 0, |c| < 1$.
- (d) $(c^n/n!)$, for $c \in \mathbb{R}$.
- (e) $(n^p/n!)$, for $p > 0$.

Proof (a) To prove that $(1/n^p)$ is null for $p > 0$, we apply the Power Rule to the sequence $(1/n)$, which we know is null.

(b) To prove that (c^n) is null for $|c| < 1$, first note that it is sufficient to consider only the case $0 \leq c < 1$, because any sequence (a_n) is null if and only if the sequence $(|a_n|)$ is null.

If $c = 0$, then the sequence is obviously null. Thus we can assume that $0 < c < 1$, so we can write

$$c = \frac{1}{1+a}, \quad \text{where } a > 0.$$

Expressing c in this way enables us to use the Binomial Theorem from Unit D1, which says that, for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$,

$$(1+x)^n = 1 + nx + \frac{1}{2}n(n-1)x^2 + \cdots + x^n.$$

Here we put $x = a$, and since a is positive, so is every term on the right-hand side.

By the Binomial Theorem

$$(1+a)^n \geq 1 + na \geq na, \quad \text{for } n = 1, 2, \dots,$$

so

$$c^n = \frac{1}{(1+a)^n} \leq \frac{1}{na}, \quad \text{for } n = 1, 2, \dots$$

Since $(1/n)$ is null, we deduce that $(1/(na))$ is null, by the Multiple Rule. Hence (c^n) is null, by the Squeeze Rule, as required.

- (c) To prove that $(n^p c^n)$ is null, for $p > 0$ and $|c| < 1$, we can again assume that $0 < c < 1$, so

$$c = \frac{1}{1+a}, \quad \text{where } a > 0.$$

First we deal with the case $p = 1$; that is, we consider the sequence (nc^n) .

 We use the Binomial Theorem again, but this time include a further term in the expansion. 

By the Binomial Theorem,

$$(1+a)^n \geq 1 + na + \frac{1}{2}n(n-1)a^2 \geq \frac{1}{2}n(n-1)a^2, \quad \text{for } n = 2, 3, \dots,$$

so

$$nc^n = \frac{n}{(1+a)^n} \leq \frac{n}{\frac{1}{2}n(n-1)a^2} = \frac{(2/a^2)}{n-1}, \quad \text{for } n = 2, 3, \dots$$

Now the sequence (a_n) defined by

$$a_n = \frac{(2/a^2)}{n-1}, \quad n = 2, 3, \dots$$

is the same as the sequence defined by

$$a_n = \frac{(2/a^2)}{n}, \quad n = 1, 2, \dots,$$

so (a_n) is null by the Multiple Rule. Hence (nc^n) is null, by the Squeeze Rule. This proves part (c) in the case $p = 1$.

To deduce that $(n^p c^n)$ is null for any $p > 0$ and $0 < c < 1$, we note that

$$n^p c^n = (nd^n)^p, \quad \text{for } n = 1, 2, \dots,$$

where $d = c^{1/p}$.

 Notice that the sequence (nd^n) is in a form that enables us to apply part (c) in the case $p = 1$, which we have just proved. 

Since $0 < d < 1$, we know that (nd^n) is null, so $(n^p c^n)$ is null for any $p > 0$, by the Power Rule.

- (d) To prove that $(c^n/n!)$ is null, we can again assume that $c > 0$. We first choose an integer m such that $m+1 > c$. Then, for $n > m+1$,

$$\begin{aligned} \frac{c^n}{n!} &= \left(\frac{c}{1}\right)\left(\frac{c}{2}\right)\cdots\left(\frac{c}{m}\right)\left(\frac{c}{m+1}\right)\cdots\left(\frac{c}{n-1}\right)\left(\frac{c}{n}\right) \\ &\leq \left(\frac{c}{1}\right)\left(\frac{c}{2}\right)\cdots\left(\frac{c}{m}\right) \times \frac{c}{n} \\ &= K \times \frac{c}{n}, \end{aligned}$$

where $K = c^m/m!$ is a constant.

 Here we have used that fact that, since $c < m+1$, it follows that $c/d < 1$ for any $d \geq m+1$. 

Since $(1/n)$ is null, we deduce that (Kc/n) is null, by the Multiple Rule. Hence $(c^n/n!)$ is null, by the Squeeze Rule.

(e) To prove that $(n^p/n!)$ is null for $p > 0$, we write

$$\frac{n^p}{n!} = \left(\frac{n^p}{2^n} \right) \left(\frac{2^n}{n!} \right), \quad \text{for } n = 1, 2, \dots$$

Since $(n^p/2^n)$ is a null sequence, by part (c) with $c = 1/2$, and $(2^n/n!)$ is also null, by part (d) with $c = 2$, we deduce that $(n^p/n!)$ is null, by the Product Rule. ■

Exercise D30

Verify that each of the following sequences (a_n) is a basic null sequence by identifying its type from those listed in Theorem D7, giving the values of c and/or p in each case.

(a) $a_n = 0.9^n, \quad n = 1, 2, \dots$

(b) $a_n = \frac{27^n}{n!}, \quad n = 1, 2, \dots$

(c) $a_n = \frac{1}{\sqrt{n}}, \quad n = 1, 2, \dots$

(d) $a_n = \frac{n^{27}}{n!}, \quad n = 1, 2, \dots$

(e) $a_n = n \left(\frac{1}{2} \right)^n, \quad n = 1, 2, \dots$

3 Convergent sequences

In the previous section we looked at null sequences, that is, sequences which converge to 0. We now turn our attention to sequences which converge to limits other than 0.

3.1 What is a convergent sequence?

The following exercise should help to give you some understanding of the behaviour of a sequence which has a limit that is not 0.

Exercise D31

Consider the sequence

$$a_n = \frac{n+1}{n}, \quad n = 1, 2, \dots$$

- (a) Draw the sequence diagram of (a_n) and describe (informally) how this sequence behaves.
- (b) What can you say (formally) about the behaviour of the sequence

$$b_n = a_n - 1, \quad n = 1, 2, \dots ?$$

The terms of the sequence (a_n) in Exercise D31 appear to approach arbitrarily close to 1; that is, the sequence (a_n) appears to converge to 1. If we subtract 1 from each term a_n to form the sequence (b_n) , then we obtain a null sequence. This example suggests the following definition of a *convergent sequence*.

Definitions

The sequence (a_n) is **convergent** with **limit l** if $(a_n - l)$ is a null sequence. We say that (a_n) **converges to l** and we write

either $\lim_{n \rightarrow \infty} a_n = l$
 or $a_n \rightarrow l$ as $n \rightarrow \infty$.

Remarks

1. The statements in the definition are read as:
 ‘the limit of a_n , as n tends to infinity, is l ';
 ‘ a_n tends to l , as n tends to infinity'.
2. Often we omit ‘as $n \rightarrow \infty$ '.
3. Do not let this use of the symbol ∞ tempt you to think that ∞ is a real number. Instead, you should remember that the phrase ‘ a_n tends to l , as n tends to infinity' means that ‘as n gets larger and larger, a_n gets closer and closer to l '.

The following are examples of convergent sequences:

- every null sequence converges to 0
- every constant sequence (c) converges to c
- as you saw in Exercise D31, the sequence (a_n) defined by

$$a_n = \frac{n+1}{n}, \quad n = 1, 2, \dots$$

is convergent and $\lim_{n \rightarrow \infty} a_n = 1$.

Exercise D32

Show that the sequence

$$a_n = \frac{n^3 + 1}{2n^3}, \quad n = 1, 2, \dots$$

converges to $\frac{1}{2}$, by considering $a_n - \frac{1}{2}$.

The definition of convergence of a sequence is often given in the following equivalent (alternative) form, mirroring the definition of a null sequence given in the previous section.

Definition (alternative)

The sequence (a_n) **converges to l** if

for each positive number ε , there is an integer N such that

$$|a_n - l| < \varepsilon, \quad \text{for all } n > N.$$

Remarks

1. In terms of the sequence diagram for (a_n) , this definition states that:
for each positive number ε , the terms a_n eventually lie inside the horizontal strip from $l - \varepsilon$ to $l + \varepsilon$.

This is illustrated in Figure 9.

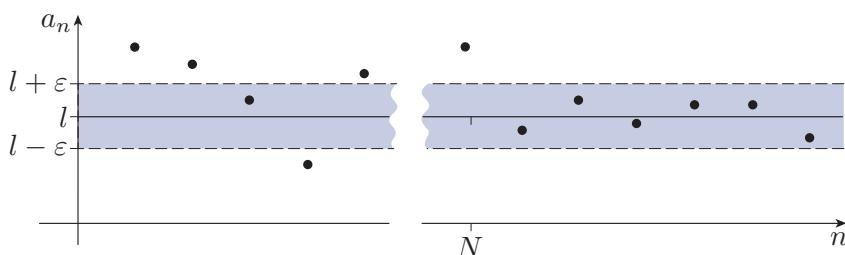


Figure 9 A sequence which converges to l

2. If a sequence is convergent, then it has a unique limit. A proof of this seemingly obvious fact is given later in this section.
3. If a given sequence converges to l , then this remains true if we add, delete or alter a finite number of terms. This follows from the corresponding result for null sequences.
4. Not all sequences are convergent, as you will see in Section 4. For example, the sequence $((-1)^n)$ is not convergent.

3.2 Combination Rules for convergent sequences

So far you have tested the convergence of a given sequence (a_n) by calculating $a_n - l$ and showing that $(a_n - l)$ is null. This presupposes that you know in advance the value of l . Usually, however, you are given a sequence (a_n) and asked to decide whether or not it converges and, if it does, to *find* its limit. Fortunately, the convergence of many sequences can be proved by using the following Combination Rules, which extend the Combination Rules for null sequences.

Theorem D8 Combination Rules for convergent sequences

If $\lim_{n \rightarrow \infty} a_n = l$ and $\lim_{n \rightarrow \infty} b_n = m$, then:

Sum Rule $\lim_{n \rightarrow \infty} (a_n + b_n) = l + m$

Multiple Rule $\lim_{n \rightarrow \infty} (\lambda a_n) = \lambda l$, for $\lambda \in \mathbb{R}$

Product Rule $\lim_{n \rightarrow \infty} (a_n b_n) = lm$

Quotient Rule $\lim_{n \rightarrow \infty} \left(\frac{a_n}{b_n} \right) = \frac{l}{m}$, provided that $m \neq 0$.

In applications of the Quotient Rule, some terms b_n can take the value 0, in which case a_n/b_n is not defined. However, we shall see (in Lemma D9) that because $m \neq 0$ this occurs for only *finitely many* b_n , so (b_n) is eventually non-zero. Thus the statement of the Quotient Rule does make sense.

We prove the Combination Rules at the end of this subsection, but first we illustrate how to apply them. When using these rules, there is no need for you to identify which particular rule you are using: you can refer simply to the Combination Rules.

Applying the Combination Rules

It is straightforward to apply the Combination Rules to simple sums, multiples, products or quotients of sequences that you already know are convergent. For example, since you know that $(1/n^2)$ is a basic null sequence, and that the constant sequence (c) has limit c , you can deduce from the Sum Rule that

$$\lim_{n \rightarrow \infty} \left(c + \frac{1}{n^2} \right) = c + 0 = c,$$

and from the Quotient Rule that

$$\lim_{n \rightarrow \infty} \left(\frac{1}{cn^2} \right) = \frac{0}{c} = 0.$$

However, the real power of the Combination Rules is that they enable us to determine the limits of some more complicated sequences, whose n th term is given by a quotient that at first sight may not seem to involve convergent sequences. This is illustrated in the next worked exercise.

Worked Exercise D25

Show that each of the following sequences (a_n) is convergent and find its limit.

$$(a) \quad a_n = \frac{(2n+1)(n+2)}{3n^2 + 3n}, \quad n = 1, 2, \dots$$

$$(b) \quad a_n = \frac{2n^2 + 10^n}{n! + 3n^3}, \quad n = 1, 2, \dots$$

Solution

Although the expressions for a_n are quotients, we cannot apply the Quotient Rule immediately because the sequences defined by the numerators and the denominators are not convergent. In each case, however, we can rearrange the expressions for a_n and then apply the Combination Rules.

(a) Dividing both the numerator and the denominator by n^2 gives

$$a_n = \frac{(2n+1)(n+2)}{3n^2 + 3n} = \frac{(2 + 1/n)(1 + 2/n)}{3 + 3/n}.$$

Since $(1/n)$ is a basic null sequence, we find, by the Combination Rules, that

$$\lim_{n \rightarrow \infty} a_n = \frac{(2 + 0)(1 + 0)}{3 + 0} = \frac{2}{3}.$$

(b) Dividing both the numerator and the denominator by $n!$ gives

$$a_n = \frac{2n^2 + 10^n}{n! + 3n^3} = \frac{2n^2/n! + 10^n/n!}{1 + 3n^3/n!}.$$

Since $(n^2/n!), (10^n/n!)$ and $(n^3/n!)$ are all basic null sequences, we find, by the Combination Rules, that (a_n) is convergent and

$$\lim_{n \rightarrow \infty} a_n = \frac{0 + 0}{1 + 0} = 0.$$

In Worked Exercise D25 the key step in determining the limit of a given sequence was to rearrange the expression for a_n in a way that enabled us to apply the Combination Rules. To explain how this is done, we need the following definition.

Definition

The **dominant term** of a quotient involving the variable n , where $n = 1, 2, \dots$, is the term in n (without its coefficient) which eventually has the largest absolute value.

As a simple example, consider the quotient

$$\frac{n^3 + 1}{2n^3}, \quad n = 1, 2, \dots,$$

which is the formula for the n th term of the sequence you met in Exercise D32. In this quotient there are two terms in n , namely n^3 and $2n^3$. To find the dominant term we exclude the coefficients, so both terms reduce to n^3 ; hence, n^3 is the dominant term in this quotient.

The method used to rearrange the expressions for a_n in Worked Exercise D25 was to divide both the numerator and the denominator by the dominant term of the quotient. Doing this converts the quotient into a form where the Combination Rules can be applied.

- In part (a) the dominant term is n^2 , which is the highest power of n in the quotient. We then used the fact that $(1/n)$ is a null sequence to find the limit of (a_n) using the Combination Rules.
- In part (b) the dominant term is $n!$, because $n!$ eventually becomes larger than n^2 , n^3 and 10^n as n increases. We then used the fact that $(n^2/n!), (10^n/n!)$ and $(n^3/n!)$ are all basic null sequences to find the limit of (a_n) using the Combination Rules.

These examples illustrate the following general strategy, where the list of dominant terms follows from the list of basic null sequences.

Strategy D7

To evaluate the limit of a sequence whose n th term is a complicated quotient, do the following.

1. Identify the dominant term, noting that

$n!$ dominates c^n ,

and, for $|c| > 1$ and $p > 0$,

c^n dominates n^p .

2. Divide both numerator and denominator by the dominant term.
3. Apply the Combination Rules.

Exercise D33

Show that each of the following sequences (a_n) is convergent and find its limit.

$$(a) \quad a_n = \frac{n^3 + 2n^2 + 3}{2n^3 + 1}, \quad n = 1, 2, \dots$$

$$(b) \quad a_n = \frac{n^2 + 2^n}{3^n + n^3}, \quad n = 1, 2, \dots$$

Hint: You can use the fact that $(2^n/3^n)$ is a basic null sequence, because $2^n/3^n = (2/3)^n$, for $n = 1, 2, \dots$.

$$(c) \quad a_n = \frac{n! + (-1)^n}{2^n + 3n!}, \quad n = 1, 2, \dots$$

Proofs of the Combination Rules

We now prove the Sum Rule, the Multiple Rule and the Product Rule by using the corresponding Combination Rules for null sequences.

Remember that $\lim_{n \rightarrow \infty} a_n = l$ means that $(a_n - l)$ is a null sequence.

Sum Rule for convergent sequences

If $\lim_{n \rightarrow \infty} a_n = l$ and $\lim_{n \rightarrow \infty} b_n = m$, then

$$\lim_{n \rightarrow \infty} (a_n + b_n) = l + m.$$

Proof We know that $(a_n - l)$ and $(b_n - m)$ are null sequences. Since

$$(a_n + b_n) - (l + m) = (a_n - l) + (b_n - m),$$

we deduce that $((a_n + b_n) - (l + m))$ is null, by the Sum Rule for null sequences. ■

We now prove the Product Rule. The Multiple Rule is a special case of the Product Rule in which the sequence (b_n) is a constant sequence.

Product Rule for convergent sequences

If $\lim_{n \rightarrow \infty} a_n = l$ and $\lim_{n \rightarrow \infty} b_n = m$, then

$$\lim_{n \rightarrow \infty} (a_n b_n) = lm.$$

Proof Here we express $a_n b_n - lm$ in terms of $a_n - l$ and $b_n - m$:

$$a_n b_n - lm = (a_n - l)(b_n - m) + m(a_n - l) + l(b_n - m).$$

Since $(a_n - l)$ and $(b_n - m)$ are null, we deduce that $(a_n b_n - lm)$ is null, by the Combination Rules for null sequences. ■

To prove the Quotient Rule we use the following lemma, which shows that if the limit of a sequence is positive, then the terms of the sequence must eventually be positive.

Lemma D9

If $\lim_{n \rightarrow \infty} a_n = l$ and $l > 0$, then there is an integer N such that

$$a_n > \frac{1}{2}l, \quad \text{for all } n > N.$$

Proof By taking $\varepsilon = \frac{1}{2}l$ in the alternative definition of convergence from Subsection 3.1, we see that there is an integer N such that

$$|a_n - l| < \frac{1}{2}l, \quad \text{for all } n > N.$$

This is illustrated in Figure 10.

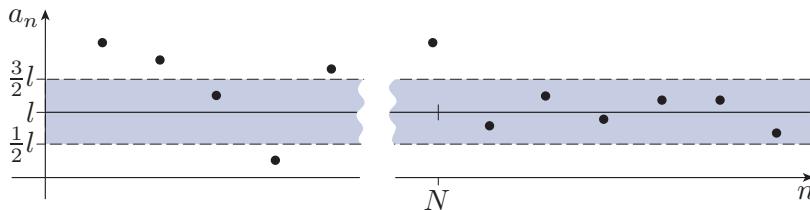


Figure 10 The sequence diagram for (a_n)

Hence

$$-\frac{1}{2}l < a_n - l < \frac{1}{2}l, \quad \text{for all } n > N,$$

and the left-hand inequality gives

$$\frac{1}{2}l < a_n, \quad \text{for all } n > N,$$

as required. ■

We now prove the Quotient Rule, which completes the proof of Theorem D8.

Quotient Rule for convergent sequences

If $\lim_{n \rightarrow \infty} a_n = l$ and $\lim_{n \rightarrow \infty} b_n = m$, then

$$\lim_{n \rightarrow \infty} \left(\frac{a_n}{b_n} \right) = \frac{l}{m}, \quad \text{provided that } m \neq 0.$$

Proof We give the proof for $m > 0$; the proof for the case $m < 0$ is similar. Once again the idea is to write the required expression in terms of $a_n - l$ and $b_n - m$:

$$\frac{a_n}{b_n} - \frac{l}{m} = \frac{m(a_n - l) - l(b_n - m)}{b_n m}.$$

Now, however, there is a slight problem: $(m(a_n - l) - l(b_n - m))$ is certainly a null sequence, but the denominator $b_n m$ is rather awkward. Some of the terms b_n can take the value 0, in which case the expression is undefined.

However, by Lemma D9, we know that for some integer N we have

$$b_n > \frac{1}{2}m, \quad \text{for all } n > N,$$

so the terms of (b_n) are eventually positive. Thus, for all $n > N$,

$$\begin{aligned} \left| \frac{a_n}{b_n} - \frac{l}{m} \right| &= \frac{|m(a_n - l) - l(b_n - m)|}{b_n m} \\ &\leq \frac{|m(a_n - l) - l(b_n - m)|}{\frac{1}{2}m^2}. \end{aligned}$$

Since the right-hand side defines a null sequence, it follows, by the Squeeze Rule for null sequences, that $\left(\frac{a_n}{b_n} - \frac{l}{m} \right)$ is null, as required. ■

3.3 Further rules for convergent sequences

There are several other theorems about convergent sequences, which are needed in later units. The first is a general version of the Squeeze Rule, illustrated in Figure 11.

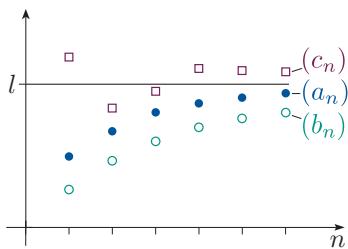


Figure 11 The Squeeze Rule

Theorem D10 Squeeze Rule for convergent sequences

If (a_n) , (b_n) and (c_n) are sequences such that

1. $b_n \leq a_n \leq c_n$, for $n = 1, 2, \dots$,
2. $\lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} c_n = l$,

then $\lim_{n \rightarrow \infty} a_n = l$.

Proof By the Combination Rules,

$$\lim_{n \rightarrow \infty} (c_n - b_n) = l - l = 0,$$

so $(c_n - b_n)$ is a null sequence. Also, by condition 1 in the statement of the theorem,

$$0 \leq a_n - b_n \leq c_n - b_n, \quad \text{for } n = 1, 2, \dots,$$

so $(a_n - b_n)$ is null, by the Squeeze Rule for null sequences.

Now we write a_n in the form

$$a_n = (a_n - b_n) + b_n.$$

Then, by the Combination Rules,

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} (a_n - b_n) + \lim_{n \rightarrow \infty} b_n = 0 + l = l.$$

Note that in applications of the Squeeze Rule, it is sufficient to check that condition 1 *eventually* holds. This is because the values of a *finite* number of terms do not affect convergence.

The following worked exercise and exercise illustrate the use of the Squeeze Rule and the Binomial Theorem in the derivation of two important limits.

Worked Exercise D26

- (a) Prove that, if $c > 0$, then

$$(1 + c)^{1/n} \leq 1 + \frac{c}{n}, \quad \text{for } n = 1, 2, \dots$$

- (b) Use the Squeeze Rule to deduce that if $a > 0$, then

$$\lim_{n \rightarrow \infty} a^{1/n} = 1.$$

Solution

- (a) We proved this inequality for the case $c = 1$ in Worked Exercise D13 of Unit D1.

Using Rule 5 for rearranging inequalities with $p = n$, we obtain

$$(1 + c)^{1/n} \leq 1 + \frac{c}{n} \iff 1 + c \leq \left(1 + \frac{c}{n}\right)^n.$$

The right-hand inequality holds because

$$\left(1 + \frac{c}{n}\right)^n \geq 1 + n \left(\frac{c}{n}\right) = 1 + c,$$

by the Binomial Theorem, so the left-hand inequality also holds.

- (b) We consider the cases $a > 1$, $a = 1$ and $0 < a < 1$ separately.

If $a > 1$, then we can write $a = 1 + c$, where $c > 0$.

In this application of the Squeeze Rule we take the ‘lower’ sequence to be the constant sequence whose terms are all equal to 1.

By part (a),

$$1 < a^{1/n} = (1 + c)^{1/n} \leq 1 + \frac{c}{n}, \quad \text{for } n = 1, 2, \dots$$

Since $(1/n)$ is a basic null sequence, it follows from the Combination Rules that $\lim_{n \rightarrow \infty} \left(1 + \frac{c}{n}\right) = 1$. We deduce, by the Squeeze Rule, that

$$\lim_{n \rightarrow \infty} a^{1/n} = 1.$$

If $a = 1$, then $a^{1/n} = 1$, for $n = 1, 2, \dots$, so

$$\lim_{n \rightarrow \infty} a^{1/n} = 1.$$

Finally, if $0 < a < 1$, then $1/a > 1$, so $\lim_{n \rightarrow \infty} (1/a)^{1/n} = 1$, by the first case. Hence, by the Quotient Rule,

$$\lim_{n \rightarrow \infty} a^{1/n} = \lim_{n \rightarrow \infty} \frac{1}{(1/a)^{1/n}} = \frac{1}{1} = 1.$$

Exercise D34

- (a) Prove that

$$n^{1/n} \leq 1 + \sqrt{\frac{2}{n-1}}, \quad \text{for } n \geq 2.$$

Hint: By the Binomial Theorem, we have

$$(1+x)^n \geq \frac{n(n-1)}{2!}x^2, \quad \text{for } n \geq 2, x \geq 0.$$

- (b) Use the Squeeze Rule to deduce from part (a) that

$$\lim_{n \rightarrow \infty} n^{1/n} = 1.$$

Next we show that taking limits preserves weak inequalities.

Theorem D11 Limit Inequality Rule

If $\lim_{n \rightarrow \infty} a_n = l$ and $\lim_{n \rightarrow \infty} b_n = m$, and also

$$a_n \leq b_n, \quad \text{for } n = 1, 2, \dots,$$

then $l \leq m$.

Proof We use proof by contradiction. Suppose that $a_n \rightarrow l$, $b_n \rightarrow m$ and $a_n \leq b_n$, for $n = 1, 2, \dots$. If $l > m$, then, by the Combination Rules,

$$\lim_{n \rightarrow \infty} (a_n - b_n) = l - m > 0.$$

Hence, by Lemma D9, there is an integer N such that

$$a_n - b_n > \frac{1}{2}(l - m), \quad \text{for all } n > N. \tag{7}$$

Since $a_n - b_n \leq 0$, for $n = 1, 2, \dots$, statement (7) gives a contradiction.

Hence the inequality $l \leq m$ is true. ■

We have the following corollary, promised earlier, that a convergent sequence has a unique limit.

Corollary D12

If $\lim_{n \rightarrow \infty} a_n = l$ and $\lim_{n \rightarrow \infty} a_n = m$, then $l = m$.

Proof Applying the Limit Inequality Rule with $b_n = a_n$, we deduce that $l \leq m$ and also that $m \leq l$. Hence $l = m$. ■

Note that taking limits does not preserve *strict* inequalities. For example, if $a_n = 1/n$, $n = 1, 2, \dots$, and $b_n = 2/n$, $n = 1, 2, \dots$, then

$$a_n < b_n, \quad \text{for } n = 1, 2, \dots$$

But it is not true that $\lim_{n \rightarrow \infty} a_n < \lim_{n \rightarrow \infty} b_n$, since both limits are 0; this is illustrated in Figure 12.

In Subsection 2.1, we pointed out that a sequence (a_n) is null if and only if the sequence $(|a_n|)$ is null. The final theorem in this section is a partial generalisation of this result.

Theorem D13

If $\lim_{n \rightarrow \infty} a_n = l$, then $\lim_{n \rightarrow \infty} |a_n| = |l|$.

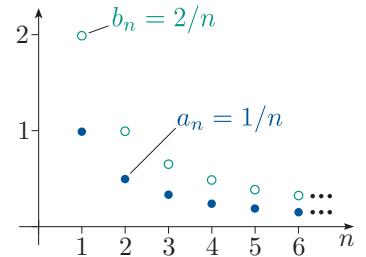


Figure 12 Two sequences with the same limit

Proof We want to show that $(|a_n| - |l|)$ is null. Using the backwards form of the Triangle Inequality, which you met in Unit D1, we obtain

$$||a_n| - |l|| \leq |a_n - l|, \quad \text{for } n = 1, 2, \dots$$

Since $(a_n - l)$ is null, so is $(|a_n - l|)$, and we deduce from the Squeeze Rule for null sequences that $(|a_n| - |l|)$ is null, as required. ■

Note that Theorem D13 is only a partial generalisation of the earlier result about null sequences because its converse is false: if $|a_n| \rightarrow |l|$, then it does not follow that $a_n \rightarrow l$. For example, consider the sequence $a_n = (-1)^n$, $n = 1, 2, \dots$; in this case,

$$|a_n| \rightarrow 1 \text{ as } n \rightarrow \infty,$$

but (a_n) does not converge.

4 Divergent sequences

In previous sections you have seen many examples of sequences that are convergent. We now investigate the behaviour of sequences which do not converge.

4.1 What is a divergent sequence?

Any sequence that does not converge is said to be *divergent*.

Definition

A sequence is **divergent** if it is not convergent.

Figure 13 gives the sequence diagrams for three different sequences (a_n) . Each of these sequences is divergent but, as you can see, they behave differently.

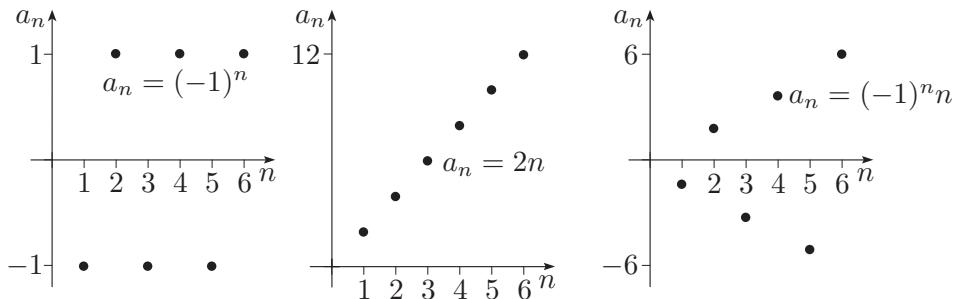


Figure 13 Three divergent sequences

It is not easy to prove from the definition that these sequences are divergent. To show that a sequence (a_n) is divergent, we would have to show that (a_n) is not convergent; that is, for *every* real number l , the sequence $(a_n - l)$ is not null.

In this section we obtain criteria for divergence which avoid us having to argue directly from the definition. At the end of the section we give a strategy for proving divergence using two criteria, which together cover all cases. We obtain these criteria by establishing certain properties which are necessarily possessed by a convergent sequence; if a sequence does *not* have one of these properties, then it must be divergent.

4.2 Bounded and unbounded sequences

One property possessed by a convergent sequence is that it must be *bounded*, as we will show.

Definitions

A sequence (a_n) is **bounded** if there is a number M such that

$$|a_n| \leq M, \quad \text{for } n = 1, 2, \dots$$

A sequence is **unbounded** if it is not bounded.

Thus a sequence (a_n) is bounded if *all* the terms a_n lie on the sequence diagram in the horizontal strip from $-M$ to M , for some positive number M ; see Figure 14.

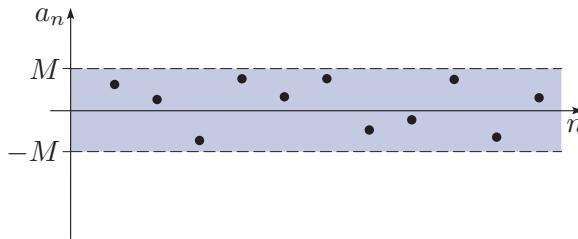


Figure 14 A bounded sequence

For example, the sequence $((-1)^n)$ is bounded because

$$|(-1)^n| \leq 1, \quad \text{for } n = 1, 2, \dots$$

However, the sequences $(2n)$ and (n^2) are unbounded since, for each positive number M , we can find terms of these sequences whose absolute values are greater than M .

Exercise D35

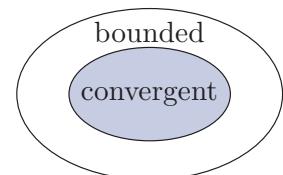
Classify the following sequences (a_n) as bounded or unbounded.

- (a) $a_n = 1 + (-1)^n, \quad n = 1, 2, \dots$
- (b) $a_n = (-1)^n n, \quad n = 1, 2, \dots$
- (c) $a_n = \frac{2n+1}{n}, \quad n = 1, 2, \dots$

The sequence $((-1)^n)$ shows that a bounded sequence is not necessarily convergent. However, we can prove that a convergent sequence is necessarily bounded. This is illustrated in Figure 15.

Theorem D14

If (a_n) is convergent, then (a_n) is bounded.



convergent \implies bounded

Figure 15 Convergent and bounded sequences

Proof We know that $a_n \rightarrow l$, for some real number l , so $(a_n - l)$ is a null sequence. Taking $\varepsilon = 1$ in the definition of a null sequence, we see that there is an integer N such that

$$|a_n - l| < 1, \quad \text{for all } n > N.$$

Now

$$\begin{aligned} |a_n| &= |(a_n - l) + l| \\ &\leq |a_n - l| + |l|, \quad \text{by the Triangle Inequality.} \end{aligned}$$

It follows that

$$|a_n| < 1 + |l|, \quad \text{for all } n > N.$$

This is the type of inequality needed to prove that (a_n) is bounded, but it does not include the terms a_1, a_2, \dots, a_N . To complete the proof, we put

$$M = \max \{|a_1|, |a_2|, \dots, |a_N|, 1 + |l|\}.$$

It then follows that

$$|a_n| \leq M, \quad \text{for } n = 1, 2, \dots,$$

as required. ■

From Theorem D14 we obtain the following test for the *divergence* of a sequence, which is illustrated in Figure 16.

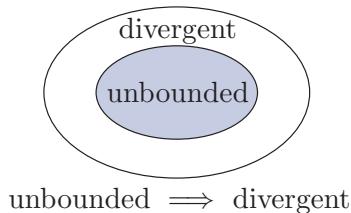


Figure 16 Unbounded and divergent sequences

Corollary D15

If (a_n) is unbounded, then (a_n) is divergent.

For example, the sequences $(2n)$ and $((-1)^n n)$ are both unbounded, so they are both divergent, by Corollary D15.

Exercise D36

Classify the following sequences (a_n) as bounded or unbounded and as convergent or divergent.

- (a) $a_n = \sqrt{n}, \quad n = 1, 2, \dots$
- (b) $a_n = \frac{n^2 + n}{n^2 + 1}, \quad n = 1, 2, \dots$
- (c) $a_n = (-1)^n n^2, \quad n = 1, 2, \dots$
- (d) $a_n = n^{(-1)^n}, \quad n = 1, 2, \dots$

4.3 Sequences tending to infinity

Although the sequences $(2n)$ and $((-1)^n n)$ are both unbounded and hence divergent, there is a marked difference in their behaviour. Informally, the terms of both sequences become arbitrarily large, but those of the sequence $(2n)$ become arbitrarily large and positive. The following definition makes this informal idea precise.

Definition

The sequence (a_n) **tends to infinity** if

for each positive number M , there is an integer N such that

$$a_n > M, \quad \text{for all } n > N.$$

In this case, we write

$$a_n \rightarrow \infty \text{ as } n \rightarrow \infty.$$

Remarks

- Often we omit ‘as $n \rightarrow \infty$ ’ and simply write $a_n \rightarrow \infty$.
- In terms of the sequence diagram for (a_n) , this definition states that, for each positive number M , the terms a_n eventually lie above the horizontal line at height M ; see Figure 17.

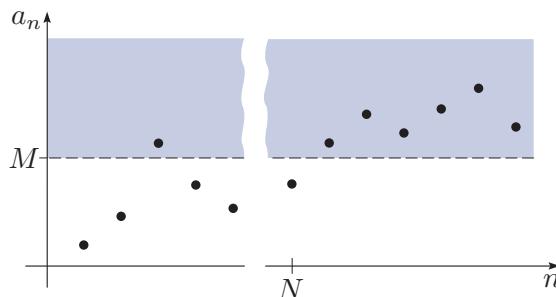


Figure 17 A sequence tending to infinity

- If a sequence tends to infinity, then it is unbounded and hence divergent, by Corollary D15.
- If a given sequence tends to infinity, then this remains true if we add, delete or alter a finite number of terms.

The next rule enables us to use our knowledge of null sequences to identify sequences which tend to infinity.

Theorem D16 Reciprocal Rule for sequences

If the sequence (a_n) satisfies the conditions

- (a_n) is eventually positive
- $(1/a_n)$ is a null sequence

then $a_n \rightarrow \infty$.

Proof To prove that $a_n \rightarrow \infty$, we have to show that:

for each positive number M , there is an integer N such that

$$a_n > M, \quad \text{for all } n > N. \quad (8)$$

Let M be a positive number. Since (a_n) is eventually positive, we can choose an integer N_1 such that

$$a_n > 0, \quad \text{for all } n > N_1.$$

Since $(1/a_n)$ is null, we can take $\varepsilon = 1/M$ in the definition of a null sequence and choose an integer N_2 such that

$$\left| \frac{1}{a_n} \right| < \frac{1}{M}, \quad \text{for all } n > N_2.$$

Now let $N = \max\{N_1, N_2\}$; then

$$0 < \frac{1}{a_n} < \frac{1}{M}, \quad \text{for all } n > N.$$

This statement is equivalent to statement (8), so $a_n \rightarrow \infty$. ■

The next worked exercise illustrates the use of the Reciprocal Rule.

Worked Exercise D27

Use the Reciprocal Rule to prove that the following sequences (a_n) tend to infinity.

- (a) $a_n = n^3/2, \quad n = 1, 2, \dots$
- (b) $a_n = n! + 10^n, \quad n = 1, 2, \dots$
- (c) $a_n = n! - 10^n, \quad n = 1, 2, \dots$

Solution

- (a) Each term of the sequence (a_n) is positive and

$$\frac{1}{a_n} = \frac{2}{n^3}.$$

Now $(1/n^3)$ is a basic null sequence, so $(2/n^3)$ is null, by the Multiple Rule.

Hence $a_n \rightarrow \infty$, by the Reciprocal Rule.

- (b) Each term of the sequence $(n! + 10^n)$ is positive.

The dominant term is $n!$, so we write

$$\frac{1}{a_n} = \frac{1}{n! + 10^n} = \frac{1/n!}{1 + 10^n/n!}.$$

Now, $(1/n!)$ and $(10^n/n!)$ are basic null sequences. Thus, by the Combination Rules,

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} = \frac{0}{1 + 0} = 0.$$

Alternatively, you could argue that, since $\frac{1}{n! + 10^n} \leq \frac{1}{n!}$, the sequence $\left(\frac{1}{n! + 10^n}\right)$ is null by the Squeeze Rule for null sequences.

It follows that $a_n \rightarrow \infty$, by the Reciprocal Rule.

- (c) The first few terms of this sequence are *not* positive but we can show that (a_n) is *eventually* positive.

The dominant term is $n!$, so we first write

$$n! - 10^n = n!(1 - 10^n/n!), \quad n = 1, 2, \dots$$

Since $(10^n/n!)$ is a basic null sequence, we know that $10^n/n!$ is eventually less than 1, so $(n! - 10^n)$ is eventually positive.

Next we write

$$\frac{1}{a_n} = \frac{1}{n! - 10^n} = \frac{1/n!}{1 - 10^n/n!}.$$

Now, $(1/n!)$ and $(10^n/n!)$ are basic null sequences. Thus, by the Combination Rules,

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} = \frac{0}{1 - 0} = 0.$$

Hence $a_n \rightarrow \infty$, by the Reciprocal Rule.

There are also versions of the Combination Rules and Squeeze Rule for sequences which tend to infinity. We state these without proof. Recall that \mathbb{R}^+ is the set of positive real numbers; that is, $\mathbb{R}^+ = \{x : x > 0\}$.

Theorem D17 Combination Rules for sequences which tend to infinity

If (a_n) tends to infinity and (b_n) tends to infinity, then

Sum Rule $(a_n + b_n)$ tends to infinity

Multiple Rule (λa_n) tends to infinity, for $\lambda \in \mathbb{R}^+$

Product Rule $(a_n b_n)$ tends to infinity.

Theorem D18 Squeeze Rule for sequences which tend to infinity

If (b_n) tends to infinity and

$$a_n \geq b_n, \quad \text{for } n = 1, 2, \dots,$$

then (a_n) tends to infinity.

Exercise D37

For each of the following sequences (a_n) , prove that $a_n \rightarrow \infty$.

- (a) $a_n = 2^n/n$, $n = 1, 2, \dots$
- (b) $a_n = 2^n - n^9$, $n = 1, 2, \dots$
- (c) $a_n = 2^n/n + 5n^9$, $n = 1, 2, \dots$
- (d) $a_n = \frac{2^n + n^2}{n^9 + n}$, $n = 1, 2, \dots$

We can also define what it means for a sequence (a_n) to *tend to minus infinity*. (Note that, in some texts, the symbol $+\infty$ is used for sequences that tend to ∞ , in order to have symmetry with the symbol $-\infty$.)

Definition

The sequence (a_n) **tends to minus infinity** if

$$-a_n \rightarrow \infty \text{ as } n \rightarrow \infty.$$

We write

$$a_n \rightarrow -\infty \text{ as } n \rightarrow \infty.$$

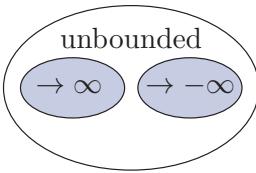


Figure 18 Unbounded sequences

For example, the sequence $(-n^2)$ tends to minus infinity because (n^2) tends to infinity. Sequences which tend to minus infinity are unbounded and hence divergent. However, the sequence $((-1)^n n)$ shows that an unbounded sequence need not tend to either infinity or to minus infinity; see Figure 18.

4.4 Subsequences

In this subsection we give two more criteria for establishing that a sequence diverges; both involve the idea of a *subsequence*. For example, consider the bounded divergent sequence $((-1)^n)$. This sequence splits naturally into two:

- the *even* terms $a_2, a_4, \dots, a_{2k}, \dots$, each of which equals 1
- the *odd* terms $a_1, a_3, \dots, a_{2k-1}, \dots$, each of which equals -1 .

Both of these are sequences in their own right, and we call them the **even subsequence** (a_{2k}) and the **odd subsequence** (a_{2k-1}) . This is illustrated in Figure 19.

In general, given a sequence (a_n) , we can consider many different subsequences, such as:

- (a_{3k}) , comprising the terms a_3, a_6, a_9, \dots
- (a_{4k+1}) , comprising the terms a_5, a_9, a_{13}, \dots
- $(a_{k!})$, comprising the terms a_1, a_2, a_6, \dots

We assume that k takes the values from 1 to ∞ unless specified otherwise.

Definition

The sequence (a_{n_k}) is a **subsequence** of the sequence (a_n) if (n_k) is a strictly increasing sequence of positive integers; that is,

$$n_1 < n_2 < n_3 < \dots$$

The subsequence (a_{n_k}) of the sequence (a_n) has terms

$$a_{n_1}, a_{n_2}, a_{n_3}, \dots,$$

and is often specified by a formula giving n_k in terms of k . For example, the subsequence $(a_{n_k}) = (a_{5k+2})$ corresponds to those terms from (a_n) whose subscripts are given by positive integers

$$n_k = 5k + 2, \quad k = 1, 2, \dots$$

Thus the first term of (a_{5k+2}) is a_7 , the second is a_{12} , and so on.

Note that if (n_k) is any strictly increasing sequence of positive integers, then $n_k \geq k$, for $k = 1, 2, \dots$, so $n_k \rightarrow \infty$ as $k \rightarrow \infty$.

Exercise D38

- (a) Let $a_n = n^2$, $n = 1, 2, \dots$. Write down the first five terms of each of the subsequences (a_{n_k}) , where:
- (i) $n_k = 2k$
 - (ii) $n_k = 4k - 1$
 - (iii) $n_k = k^2$.

- (b) Write down the first three terms of the odd and even subsequences of the sequence (a_n) defined by

$$a_n = n^{(-1)^n}, \quad n = 1, 2, \dots$$

Next we show that certain properties of sequences are inherited by their subsequences.

Theorem D19

For any subsequence (a_{n_k}) of a sequence (a_n) :

- (a) if $a_n \rightarrow l$ as $n \rightarrow \infty$, then $a_{n_k} \rightarrow l$ as $k \rightarrow \infty$
- (b) if $a_n \rightarrow \infty$ as $n \rightarrow \infty$, then $a_{n_k} \rightarrow \infty$ as $k \rightarrow \infty$.

Proof We prove only part (a); the proof of part (b) is similar. We want to show that

for each positive number ε , there is a positive integer K such that

$$|a_{n_k} - l| < \varepsilon, \quad \text{for all } k > K. \quad (9)$$

Let ε be a positive number. Since $(a_n - l)$ is null, we know that there is a positive integer N such that

$$|a_n - l| < \varepsilon, \quad \text{for all } n > N.$$

If we take K so large that $n_K \geq N$, then

$$n_k > n_K \geq N, \quad \text{for all } k > K.$$

Hence statement (9) holds for this value of K , as required. ■

The following criteria for establishing that a sequence is *divergent* are immediate consequences of Theorem D19(a).

Corollary D20 Subsequence Rules

First Subsequence Rule The sequence (a_n) is divergent if (a_n) has two convergent subsequences with different limits.

Second Subsequence Rule The sequence (a_n) is divergent if (a_n) has a subsequence which tends to infinity or a subsequence which tends to minus infinity.

We can now formulate a general strategy for showing that a sequence is divergent, as promised at the beginning of this section.

Strategy D8

To prove that the sequence (a_n) is divergent, either:

- show that (a_n) has two convergent subsequences with different limits, or
- show that (a_n) has a subsequence which tends to infinity or a subsequence which tends to minus infinity.

For example, the sequence $((-1)^n)$ has two convergent subsequences which have different limits, namely, the even subsequence with limit 1 and the odd subsequence with limit -1 . So the sequence $((-1)^n)$ is divergent, by the First Subsequence Rule.

On the other hand, the sequence $(n^{(-1)^n})$ has a subsequence (the even subsequence) which tends to infinity since, if $n = 2k$, then $n^{(-1)^n} = 2k$. So $(n^{(-1)^n})$ is divergent, by the Second Subsequence Rule.

To apply Strategy D8 successfully, you need to be able to recognise convergent subsequences with different limits, or a subsequence (which may be the whole sequence) which tends to infinity or to minus infinity. It is not always easy to do this, and some experimentation may be required. If the formula for a_n involves the expression $(-1)^n$, it is a good idea to consider the odd and even subsequences, although this may not always work. It may be helpful to calculate the values of the first few terms in order to try to identify suitable subsequences.

Exercise D39

Use Strategy D8 to prove that each of the following sequences (a_n) is divergent. Remember that $\lfloor x \rfloor$ denotes the integer part of x .

$$(a) \quad a_n = (-1)^n + \frac{1}{n}, \quad n = 1, 2, \dots$$

$$(b) \quad a_n = \frac{1}{3}n - \lfloor \frac{1}{3}n \rfloor, \quad n = 1, 2, \dots$$

$$(c) \quad a_n = n \sin\left(\frac{1}{2}n\pi\right), \quad n = 1, 2, \dots$$

We end this section by giving a result about subsequences which will be needed in later analysis units.

We say that a sequence (a_n) *consists* of two subsequences (a_{m_k}) and (a_{n_k}) when every term of the sequence appears in one or other of the subsequences. For example, every sequence (a_n) consists of its even subsequence (a_{2k}) and its odd subsequence (a_{2k-1}) . The next theorem tells us that, in these circumstances, if the two subsequences tend to the same limit, then so does the whole sequence.

Theorem D21

Let (a_n) consist of two subsequences (a_{m_k}) and (a_{n_k}) , which both tend to the *same* limit l . Then

$$\lim_{n \rightarrow \infty} a_n = l.$$

Proof We want to show that

for each $\varepsilon > 0$, there is an integer N such that

$$|a_n - l| < \varepsilon, \quad \text{for all } n > N. \quad (10)$$

Let ε be a positive number. We know that there are integers K_1 and K_2 such that

$$|a_{m_k} - l| < \varepsilon, \quad \text{for all } k > K_1,$$

and

$$|a_{n_k} - l| < \varepsilon, \quad \text{for all } k > K_2.$$

Now let

$$N = \max\{m_{K_1}, n_{K_2}\}.$$

Since each $n > N$ is either of the form m_k , with $k > K_1$, or of the form n_k , with $k > K_2$, we deduce that statement (10) holds with this value of N . ■

5 Monotone Convergence Theorem

In this section you will see a proof of the Monotone Convergence Theorem, which states that any increasing sequence which is bounded above must be convergent.

We illustrate the theorem with particular sequences that converge to π and to e . These applications of the Monotone Convergence Theorem are not assessed, but you should make sure that you have a good understanding of the statement of the theorem as we will use it in later analysis units.

5.1 Convergence of monotonic sequences

In Section 3 you met various techniques for finding the limit of a convergent sequence. As a result, you may be under the impression that, if we know that a sequence converges, then we can always find its limit. However, it is sometimes possible to prove that a sequence is convergent without being able to find its limit.

For example, this situation can occur with a given sequence (a_n) with the following two properties:

1. (a_n) is an *increasing* sequence
2. (a_n) is *bounded above*; that is, there is a real number M such that

$$a_n \leq M, \quad \text{for } n = 1, 2, \dots$$

We will prove that such a sequence must be convergent. Likewise, if (a_n) is a sequence which is *decreasing* and *bounded below*, then (a_n) must be convergent. These ideas are illustrated in Figure 20.

The next theorem combines these results.

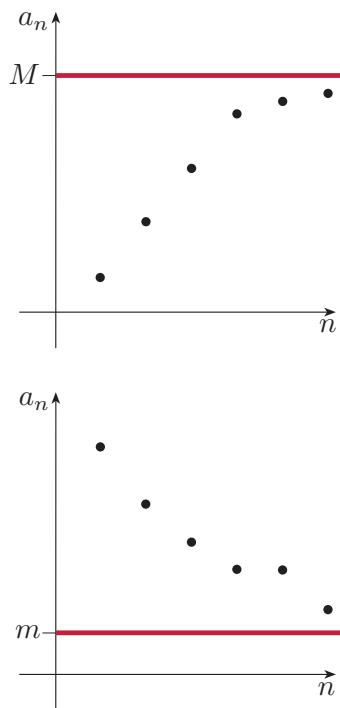


Figure 20 An increasing sequence that is bounded above and a decreasing sequence that is bounded below

Theorem D22 Monotone Convergence Theorem

If the sequence (a_n) is either

- increasing and bounded above, or
- decreasing and bounded below,

then (a_n) is convergent.

Proof We prove only that (a_n) is convergent if it is increasing and bounded above; the proof where (a_n) is decreasing and bounded below is similar.

Since (a_n) is bounded above, the set $\{a_n : n = 1, 2, \dots\}$ has a least upper bound, l say.

This follows from the Least Upper Bound Property of \mathbb{R} , which you met in Subsection 4.3 of Unit D1. The sequence (a_n) is illustrated in Figure 21.

We now prove that

$$\lim_{n \rightarrow \infty} a_n = l.$$

We want to show that

for each $\varepsilon > 0$, there is an integer N such that

$$|a_n - l| < \varepsilon, \quad \text{for all } n > N. \quad (11)$$

Let ε be a positive number. Since l is the least upper bound of the set $\{a_n : n = 1, 2, \dots\}$, there is an integer N such that

$$a_N > l - \varepsilon.$$

 If this were not true, then $l - \varepsilon$ would be an upper bound of the set $\{a_n : n = 1, 2, \dots\}$, contradicting the fact that l is the *least* upper bound. 

Because (a_n) is increasing, we have $a_n \geq a_N$ for $n > N$, so

$$a_n > l - \varepsilon, \quad \text{for all } n > N.$$

But then, since $a_n \leq l$ for all n , it follows that

$$|a_n - l| = l - a_n < \varepsilon, \quad \text{for all } n > N,$$

which proves statement (11). Hence (a_n) converges to l . ■

The Monotone Convergence Theorem tells us that a sequence such as $(1 - 1/n)$, which is increasing and bounded above (by 1, for example), must be convergent. In this case, of course, we already know that $(1 - 1/n)$ is convergent with limit 1, without using the Monotone Convergence Theorem.

The Monotone Convergence Theorem is most useful when we suspect that a sequence is convergent, but we cannot find the limit directly. It can also be used to give precise definitions of numbers, such as π , about which we have only an informal idea, as you will see in the next subsection.

For completeness, we point out that if (a_n) is increasing but is *not* bounded above, then $a_n \rightarrow \infty$. For if (a_n) is not bounded above then, for any real number M , we can find an integer N such that $a_N > M$. Since (a_n) is increasing, we have $a_n \geq a_N$ for $n > N$, so

$$a_n > M, \quad \text{for all } n > N.$$

Hence $a_n \rightarrow \infty$ as $n \rightarrow \infty$.

Similarly, if (a_n) is decreasing but is not bounded below, then $a_n \rightarrow -\infty$.

We now summarise all these results about monotonic sequences.

Theorem D23 Monotonic Sequence Theorem

If the sequence (a_n) is monotonic, then either (a_n) is convergent or $a_n \rightarrow \pm\infty$.

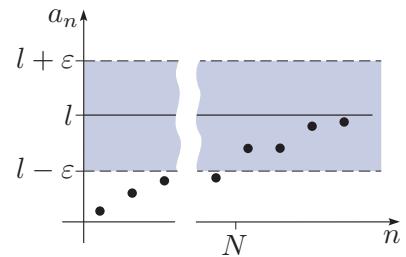


Figure 21 An increasing sequence with least upper bound l

Exercise D40

Prove that if the sequence (a_n) is increasing and has a subsequence (a_{n_k}) which is convergent, then (a_n) is convergent.

5.2 The number π

The rest of this section is not assessed but is included for your interest.

One of the oldest mathematical problems is to determine the area and the length of the perimeter of a disc of radius r . It is well known that these quantities are given by the formulas πr^2 and $2\pi r$, respectively. But what exactly is π ?

We define π by giving a precise definition of the area of a disc of radius 1. Our definition is based on a method used by Archimedes to approximate a circle of radius 1 by regular polygons circumscribed and inscribed in the circle.

Archimedes (c.287–c.212 BCE) of Syracuse was one of the greatest scientists of classical antiquity. Although many details of his life have survived, they are largely anecdotal and should be treated with caution. He had an interest in many areas of mathematics: geometry, arithmetic, astronomy, mechanics, statics (levers and centres of gravity), and hydrostatics (bodies floating in water). In particular, he anticipated modern calculus and analysis by applying the Eudoxean method of exhaustion to derive and rigorously prove a range of geometrical results including the area of a circle, the surface area and volume of a sphere and the area under a parabola. He is also credited with the design of many mechanical inventions including several war machines for use against the Roman armies laying siege to Syracuse.

What is especially noticeable about Archimedes, by comparison with many earlier mathematicians, is the way he combined pure geometrical analysis, and the mechanical or practical: his work on the lever is purely geometrical, whereas his astronomical book *On Sphere-making* (now lost) was about constructing a planetarium that modelled the motions of heavenly bodies. Indeed, his reputation in the centuries after his death was more as a maker of mechanical marvels than as a geometer.

Archimedes established bounds for the value of π by taking a circle of radius 1 and considering the perimeters of circumscribed and inscribed polygons, starting with a regular hexagon (6 sides) and progressively doubling the number of sides so as to get regular polygons of 12, 24, 48 and 96 sides, becoming ever closer to the circle. Using this method he calculated the value of π to lie between $3\frac{10}{71}$ and $3\frac{1}{7}$.

Here we will use the areas of the polygons used by Archimedes instead of their perimeters, but the details are similar. The areas of the inscribed polygons give a lower estimate for the area of the disc and hence for π . By doubling the number of sides of the polygon, we improve the estimate. The areas of the polygons can be calculated quite simply, as illustrated in Figure 22.

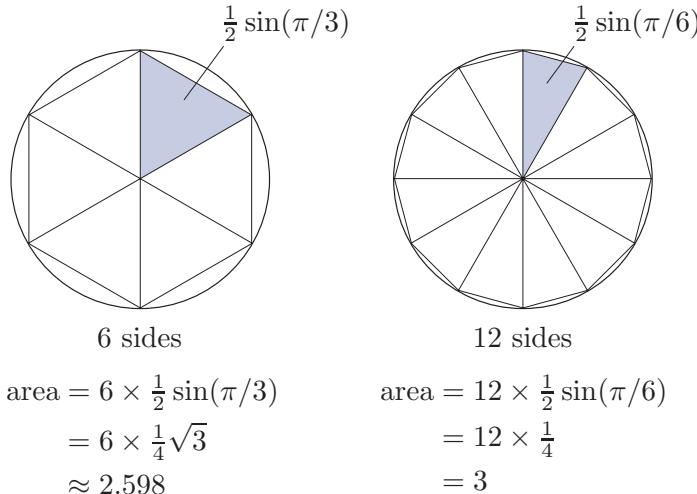


Figure 22 The area of regular inner polygons

Let s_n denote the number of sides of the n th such inner polygon, so $s_1 = 6$, $s_2 = 12$ and, in general, $s_n = 3 \times 2^n$. The n th inner polygon consists of s_n isosceles triangles, each with two equal sides of length 1 that meet at an angle $2\pi/s_n$. Thus the total area a_n of the polygon is given by

$$a_n = \frac{1}{2}s_n \sin(2\pi/s_n), \quad \text{for } n = 1, 2, \dots \quad (12)$$

For example (to 3 d.p.) we have

$$a_1 = 2.598, \quad a_2 = 3, \quad \dots, \quad a_6 = 3.141.$$

Geometrically, it is clear that each time we double the number of sides of the inner polygon, the area increases, so

$$a_1 < a_2 < a_3 < \dots < a_n < a_{n+1} < \dots$$

Hence the sequence (a_n) is (strictly) increasing.

Note that each of the polygons lies inside a square of side 2, which has area 4; see Figure 23. This implies that

$$a_n \leq 4, \quad \text{for } n = 1, 2, \dots$$

Thus the sequence (a_n) is bounded above by 4.

Hence, by the Monotone Convergence Theorem, the sequence (a_n) is convergent, with limit at most 4. Our intuitive idea of the area of the disc suggests that it is greater than each of the areas a_n , but ‘only just’. We know that the area of the circle in which the polygons are inscribed is equal to π , since it has radius 1. This leads us to make the following definition.

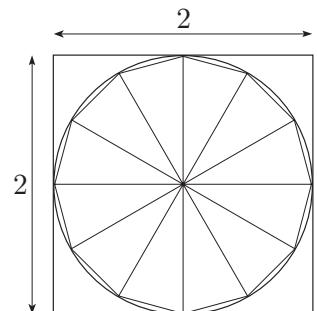


Figure 23 A circle inscribed in a square

Definition

$$\pi = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \frac{1}{2}s_n \sin(2\pi/s_n),$$

where $s_n = 3 \times 2^n$.

We will explain in a moment how to calculate the terms a_n without assuming a value for π .

First, however, we describe how to estimate the area of the disc using *outer* polygons. These give an upper estimate for π . Once again we start with a regular hexagon and repeatedly double the number of sides. The method of calculating the areas of these polygons is illustrated in Figure 24 and the explanation is given below.

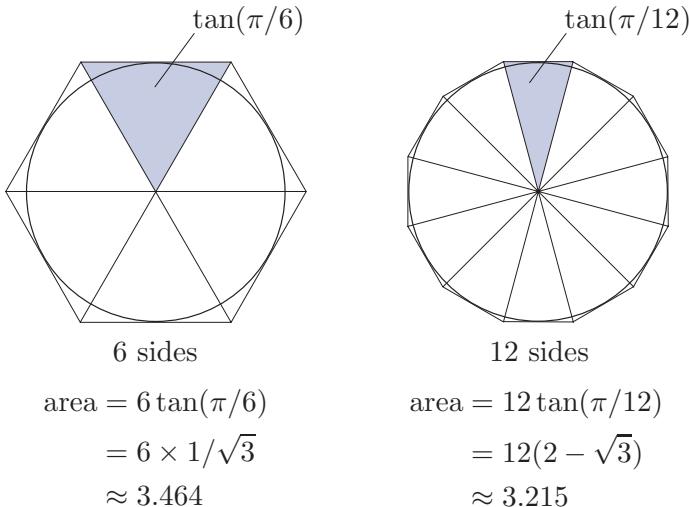


Figure 24 The area of regular outer polygons

Let s_n denote the number of sides of the n th outer polygon. As before, $s_n = 3 \times 2^n$, for $n = 1, 2, \dots$. Also let b_n denote the area of the n th outer polygon. This n th outer polygon consists of s_n isosceles triangles, each of height 1 and base $2 \tan(\pi/s_n)$. Thus

$$b_n = s_n \tan(\pi/s_n), \quad n = 1, 2, \dots \tag{13}$$

For example (to 3 d.p.) we have

$$b_1 = 3.464, \quad b_2 = 3.215, \quad \dots, \quad b_6 = 3.142.$$

Geometrically, it is clear that each time we double the number of sides of the outer polygon, the area decreases, so

$$b_1 > b_2 > b_3 > \dots > b_n > b_{n+1} > \dots$$

Hence the sequence (b_n) is (strictly) decreasing and bounded below (by 0, for example). Thus, by the Monotone Convergence Theorem, (b_n) is also convergent. Intuitively, we expect that (b_n) has the same limit as (a_n) , which we have defined to be π . But how can we prove this?

It is a remarkable fact that the terms a_n and b_n can be calculated by using the following equations, known as the *Archimedean algorithm*:

$$a_{n+1} = \sqrt{a_n b_n}, \quad n = 1, 2, \dots, \quad (14)$$

$$b_{n+1} = \frac{2a_{n+1}b_n}{a_{n+1} + b_n}, \quad n = 1, 2, \dots. \quad (15)$$

These equations for calculating a_n and b_n can be deduced from equations (12) and (13) by using trigonometric identities, though we do not prove this here.

Starting with $a_1 = \frac{3}{2}\sqrt{3} = 2.598\dots$ and $b_1 = 2\sqrt{3} = 3.464\dots$, we can use these equations iteratively to calculate first $a_2 = \sqrt{a_1 b_1}$, then

$b_2 = 2a_2 b_1 / (a_2 + b_1)$, and so on. Here are the first few values (to three decimal places) of each sequence obtained in this way.

s_n	6	12	24	48	96	192
a_n	2.598	3	3.106	3.133	3.139	3.141
b_n	3.464	3.215	3.160	3.146	3.143	3.142

It appears that the sequence (b_n) converges to the same limit as (a_n) . Indeed, by equation (14), we have $b_n = a_{n+1}^2/a_n$, so

$$\lim_{n \rightarrow \infty} b_n = \frac{\left(\lim_{n \rightarrow \infty} a_{n+1}\right)^2}{\lim_{n \rightarrow \infty} a_n} = \frac{\pi^2}{\pi} = \pi,$$

by the Combination Rules and our definition of π .

However, the convergence of these sequences (a_n) and (b_n) to $\pi = 3.14159\dots$ seems quite slow. In Unit F4 *Power series* we give other ways to calculate π , which are more efficient, and we show that π is an irrational number. All these methods use the Monotone Convergence Theorem in some way.

5.3 The number e

You will have seen that the number e plays an important role in mathematics. There are various ways in which e is defined, one of which is as the limit of the sequence (a_n) where

$$a_n = \left(1 + \frac{1}{n}\right)^n, \quad n = 1, 2, \dots.$$



Thomas Harriot

Early in the 17th century, the English mathematician Thomas Harriot (1560–1621), while working on the problem of compound interest, recognised that the sequence $a_n = (1 + 1/n)^n$ had a limit but did not give it a value. Since he did not publish his work, his results remained unknown until the much later study of his manuscripts. The first attempt to find a value for the limit of the sequence was in 1683 by Jacob Bernoulli (1654–1705), who was also working on the problem of compound interest and who calculated the limit to lie between 2 and 3. In 1748 Leonhard Euler, in his *Analysis Infinitorum*, showed that the limit is e and he calculated its approximate value to 18 decimal places.

In this subsection we use the Monotone Convergence Theorem to prove that the limit of the sequence (a_n) exists. To do this, we prove that the sequence (a_n) is increasing and bounded above. If we plot the first few terms on a sequence diagram, then it certainly seems that these properties hold; see Figure 25.

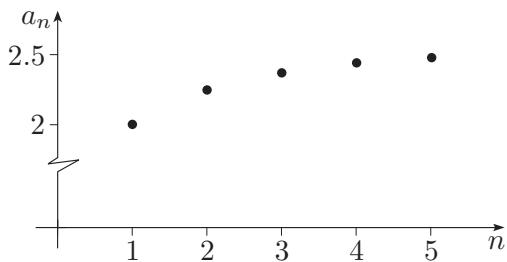


Figure 25 The sequence $a_n = (1 + \frac{1}{n})^n$

We prove these facts by using the Binomial Theorem:

$$a_n = \left(1 + \frac{1}{n}\right)^n = 1 + n\left(\frac{1}{n}\right) + \frac{n(n-1)}{2!}\left(\frac{1}{n}\right)^2 + \cdots + \left(\frac{1}{n}\right)^n.$$

As n increases, the number of terms in this sum increases and the new terms are all positive. Also, for each fixed $k \geq 1$ and any $n \geq k$, the $(k+1)$ th term of the sum is

$$\frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{1}{n}\right)^k = \frac{1}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right),$$

and the product on the right increases as n increases (because each of the factors does). Hence the sequence (a_n) is increasing.

To see that this sequence is bounded above, note that the $(k+1)$ th term of the above sum satisfies the inequality

$$\frac{1}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right) \leq \frac{1}{k!},$$

since each of the expressions in brackets is at most 1. Hence

$$\begin{aligned} a_n &= \left(1 + \frac{1}{n}\right)^n \leq 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} \\ &\leq 1 + 1 + \frac{1}{2^1} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}}, \end{aligned}$$

since $k! = k(k-1) \times \cdots \times 2 \times 1 \geq 2^{k-1}$, for $k = 1, 2, \dots$.

Now we use the fact that the sum of a finite geometric series is given by

$$1 + r + \cdots + r^{n-1} = \frac{1 - r^n}{1 - r}$$

and so, in the case $r = \frac{1}{2}$, we have

$$1 + \frac{1}{2^1} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} = 2 - \frac{1}{2^{n-1}}.$$

Thus

$$a_n = \left(1 + \frac{1}{n}\right)^n \leq 3 - \frac{1}{2^{n-1}}, \quad \text{for } n = 1, 2, \dots$$

so (a_n) is bounded above by 3.

Hence, by the Monotone Convergence Theorem, the sequence (a_n) is convergent, with limit at most 3. This allows us to make the following definition.

Definition

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

For larger and larger values of n , the terms $a_n = (1 + 1/n)^n$ give better and better approximate values for e . However, the sequence (a_n) converges to e rather slowly, and we need to take very large integers n to get a reasonable approximation to $e = 2.71828\dots$. For example,

$$\left(1 + \frac{1}{1000}\right)^{1000} = 2.716\dots$$

In Unit D3 *Series* we give another way to calculate e , which is more efficient, and we show that e is an irrational number.

Similar arguments to the ones given here can be used to prove that

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n,$$

for any $x > 0$. In fact, this is true for all real values of x but different arguments are needed when x is not positive.

Summary

In this unit you have studied sequences of real numbers denoted by (a_n) , and looked at how the ideas of convergence and the limit of a sequence can be made precise. You have seen that a key role is played by null sequences, that is, sequences which converge to zero, and met a list of basic null sequences. You have learnt how to use these basic null sequences, together with the Combination Rules and the Squeeze Rule, to show that other sequences are convergent and to find their limits.

You have also seen that many sequences are divergent. These include sequences that tend to infinity or to minus infinity, which can be identified by using the Reciprocal Rule. More generally, you have seen that a sequence is divergent if it has a subsequence that tends to infinity or minus infinity, or if it has two subsequences which converge to different limits.

Finally, you met the Monotone Convergence Theorem, which states that any increasing sequence which is bounded above must be convergent. You have also seen this theorem applied to show that particular sequences converge to the numbers π and e .

Sequences play a key role in the remaining analysis units of this module, so it is important that you have a good understanding of the material in this unit.

Learning outcomes

After working through this unit, you should be able to:

- draw the *sequence diagram* of a given sequence
- explain what is meant by a *monotonic* sequence
- explain the meaning of the phrase ‘a sequence *eventually* has a given property’
- explain the definition of *null sequence* and apply it in simple cases
- use the Power Rule, the Combination Rules and the Squeeze Rule to test for null sequences
- recognise certain *basic* null sequences
- explain what is meant by the terms *convergent sequence* and *limit of a sequence*, and by the statements $\lim_{n \rightarrow \infty} a_n = l$, or $a_n \rightarrow l$ as $n \rightarrow \infty$
- use the Combination Rules to calculate limits of sequences
- state and use some theorems about convergent sequences
- explain the terms *divergent sequence*, *bounded sequence* and *unbounded sequence*
- explain the phrases (a_n) *tends to infinity* and (a_n) *tends to minus infinity*, and use the Reciprocal Rule to recognise sequences which tend to infinity

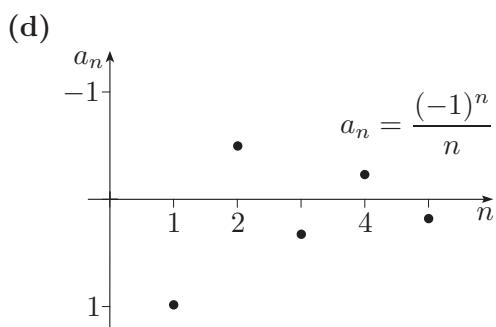
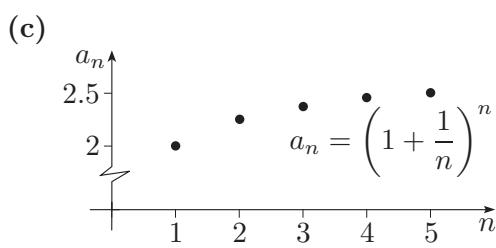
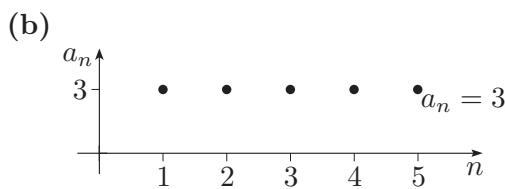
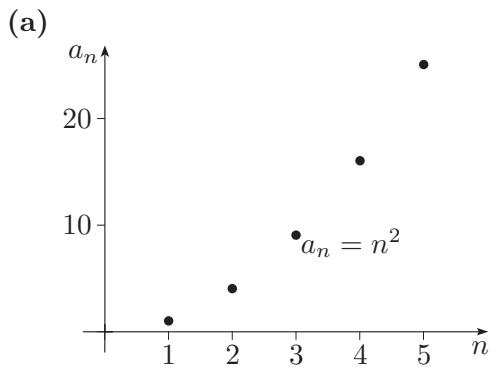
- use the Subsequence Rules to recognise divergent sequences
- state the Monotone Convergence Theorem
- understand the role of the Monotone Convergence Theorem in the definitions of the numbers π and e .

Solutions to exercises

Solution to Exercise D22

- (a) 4, 7, 10, 13, 16.
- (b) $\frac{1}{3}, \frac{1}{9}, \frac{1}{27}, \frac{1}{81}, \frac{1}{243}$.
- (c) -1, 2, -3, 4, -5.
- (d) 1, 2, 6, 24, 120.
- (e) 2, 2.25, 2.37, 2.44, 2.49.

Solution to Exercise D23



Solution to Exercise D24

- (a) Since $a_n > 0$ for all n , we can use Strategy D4. We have

$$a_n = (n-1)! \quad \text{and} \quad a_{n+1} = n!,$$

so

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{n!}{(n-1)!} = \frac{n \times (n-1)!}{(n-1)!} \\ &= n \geq 1, \quad \text{for } n = 1, 2, \dots \end{aligned}$$

Thus (a_n) is increasing, so (a_n) is monotonic. (Notice, however, that (a_n) is not strictly increasing, since $a_1 = a_2 = 1$.)

- (b) Since $a_n > 0$ for all n , we can use Strategy D4. We have

$$a_n = 2^{-n} \quad \text{and} \quad a_{n+1} = 2^{-(n+1)},$$

so

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{2^n}{2^{n+1}} = \frac{2^n}{2 \times 2^n} \\ &= \frac{1}{2} < 1, \quad \text{for } n = 1, 2, \dots \end{aligned}$$

Thus (a_n) is strictly decreasing, so (a_n) is monotonic.

- (c) We use Strategy D3. We have

$$a_n = n + \frac{1}{n} \quad \text{and} \quad a_{n+1} = n + 1 + \frac{1}{n+1},$$

so

$$\begin{aligned} a_{n+1} - a_n &= \left(n + 1 + \frac{1}{n+1}\right) - \left(n + \frac{1}{n}\right) \\ &= 1 - \frac{1}{n(n+1)} > 0, \quad \text{for } n = 1, 2, \dots \end{aligned}$$

Thus (a_n) is strictly increasing, so (a_n) is monotonic.

Solution to Exercise D25

- (a) True: $2^n > 1000$, for $n > 9$, since (2^n) is increasing and $2^{10} = 1024$.

- (b) False: all the terms a_1, a_3, a_5, \dots are negative since

$$(-1)^n = -1, \quad \text{for } n = 1, 3, 5, \dots$$

(c) True: $\frac{1}{n} < 0.025$, for $n > \frac{1}{0.025} = 40$.

(d) True: $a_n > 0$ for all n , and

$$\frac{a_{n+1}}{a_n} = \frac{1}{4} \left(\frac{n+1}{n} \right)^4.$$

Using the rules for rearranging inequalities, we have

$$\begin{aligned} \frac{1}{4} \left(\frac{n+1}{n} \right)^4 \leq 1 &\iff \left(\frac{n+1}{n} \right)^4 \leq 4 \\ &\iff 1 + \frac{1}{n} \leq 4^{1/4} \\ &\iff \frac{1}{n} \leq \sqrt[4]{2} - 1 \\ &\iff n \geq \frac{1}{\sqrt[4]{2} - 1} \approx 2.414. \end{aligned}$$

So

$$\frac{a_{n+1}}{a_n} \leq 1, \quad \text{for } n > 2.$$

Hence

$$a_{n+1} \leq a_n, \quad \text{for } n > 2,$$

so (a_n) is eventually decreasing.

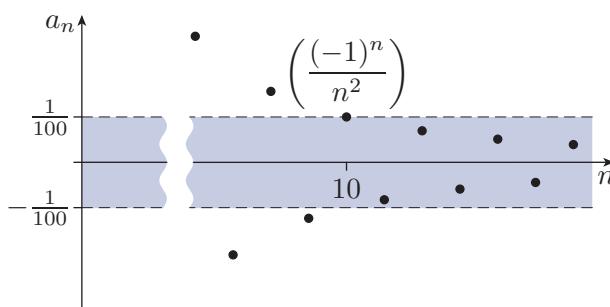
Solution to Exercise D26

Sequence diagrams are given here to aid your understanding, but you are not expected to have drawn these as part of your solutions.

(a) We have that

$$\begin{aligned} \left| \frac{(-1)^n}{n^2} \right| < \frac{1}{100} &\iff \frac{1}{n^2} < \frac{1}{100} \\ &\iff n^2 > 100 \\ &\iff n > 10. \end{aligned}$$

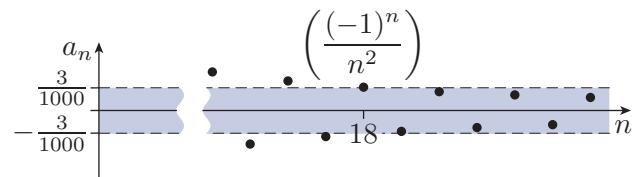
Hence we may take $N = 10$. This is illustrated below.



(b) We have that

$$\begin{aligned} \left| \frac{(-1)^n}{n^2} \right| < \frac{3}{1000} &\iff \frac{1}{n^2} < \frac{3}{1000} \\ &\iff n^2 > \frac{1000}{3} \\ &\iff n > \sqrt{\frac{1000}{3}} \approx 18.26. \end{aligned}$$

Hence we may take $N = 18$. This is illustrated below.



Solution to Exercise D27

(a) The sequence (a_n) is null. To prove this, we want to show that:

$$\text{for each } \varepsilon > 0, \text{ there is an integer } N \text{ such that} \\ \frac{1}{2n-1} < \varepsilon, \quad \text{for all } n > N. \quad (*)$$

We know that

$$\begin{aligned} \frac{1}{2n-1} < \varepsilon &\iff 2n-1 > \frac{1}{\varepsilon} \\ &\iff n > \frac{1}{2} \left(1 + \frac{1}{\varepsilon} \right), \end{aligned}$$

so statement (*) holds if we take $N \geq \frac{1}{2}(1 + 1/\varepsilon)$.

Hence (a_n) is null.

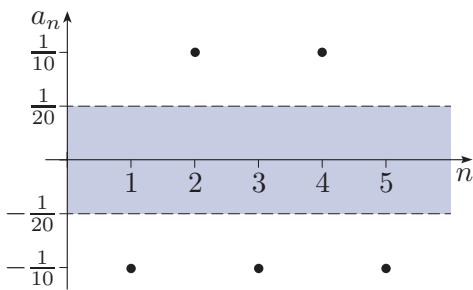
(b) The sequence (a_n) is not null. To prove this, we must find a positive value of ε for which there is no integer N such that

$$|a_n| < \varepsilon, \quad \text{for all } n > N.$$

Since

$$|a_n| = \left| \frac{(-1)^n}{10} \right| = \frac{1}{10}, \quad \text{for } n = 1, 2, \dots,$$

we can take $\varepsilon = \frac{1}{20}$. This is illustrated in the following diagram. In this particular case, *all* the terms of the sequence lie outside the ε -strip for this value of ε .



(c) The sequence (a_n) is null. To prove this, we want to show that:

for each $\varepsilon > 0$, there is an integer N such that

$$\left| \frac{(-1)^n}{n^4 + 1} \right| < \varepsilon, \quad \text{for all } n > N. \quad (*)$$

We know that

$$\left| \frac{(-1)^n}{n^4 + 1} \right| = \frac{1}{n^4 + 1}, \quad \text{for } n = 1, 2, \dots,$$

and

$$\begin{aligned} \frac{1}{n^4 + 1} < \varepsilon &\iff n^4 + 1 > \frac{1}{\varepsilon} \\ &\iff n^4 > \frac{1}{\varepsilon} - 1. \end{aligned}$$

Now $1/\varepsilon - 1$ is sometimes positive and sometimes negative, so we need to consider two cases.

If $\varepsilon \geq 1$, then $1/\varepsilon - 1 \leq 0$, so

$$n^4 > \frac{1}{\varepsilon} - 1, \quad \text{for } n = 1, 2, \dots$$

Hence statement $(*)$ holds with $N = 1$.

If $0 < \varepsilon < 1$, then $1/\varepsilon - 1 > 0$, so we can use Rule 5 for rearranging inequalities, giving

$$n^4 > \frac{1}{\varepsilon} - 1 \iff n > \left(\frac{1}{\varepsilon} - 1 \right)^{1/4}.$$

Hence statement $(*)$ holds if we take

$$N \geq (1/\varepsilon - 1)^{1/4}.$$

Thus statement $(*)$ holds in either case, so (a_n) is null.

Solution to Exercise D28

(a) We know that the sequence $\left(\frac{1}{2n-1} \right)$ is null, so (a_n) is null, by the Power Rule.

(b) We know that the sequence $\left(\frac{1}{n^3} \right)$ is null, so (a_n) is null, by the Multiple Rule.

(c) The sequences $\left(\frac{1}{n} \right)$ and $\left(\frac{1}{2n-1} \right)$ are null, so the sequences $\left(\frac{1}{n^4} \right)$ and $\left(\frac{1}{(2n-1)^{1/3}} \right)$ are also null, by the Power Rule.

Hence (a_n) is null, by the Product Rule and the Multiple Rule.

Solution to Exercise D29

(a) We guess that (a_n) is dominated by (b_n) , where

$$b_n = \frac{1}{n}, \quad n = 1, 2, \dots$$

To check this, we have to show that

$$\frac{1}{n^2 + n} \leq \frac{1}{n}, \quad \text{for } n = 1, 2, \dots$$

This holds because

$$n^2 + n \geq n, \quad \text{for } n = 1, 2, \dots$$

Since (b_n) is null, we deduce that (a_n) is null, by the Squeeze Rule.

(b) We guess that (a_n) is dominated by (b_n) , where

$$b_n = \frac{1}{n}, \quad n = 1, 2, \dots$$

To check this, we have to show that

$$\left| \frac{(-1)^n}{n!} \right| \leq \frac{1}{n}, \quad \text{for } n = 1, 2, \dots$$

This holds because

$$\left| \frac{(-1)^n}{n!} \right| = \frac{1}{n!}$$

and

$$n! \geq n, \quad \text{for } n = 1, 2, \dots$$

Since (b_n) is null, we deduce that (a_n) is null, by the Squeeze Rule.

(c) We guess that (a_n) is dominated by (b_n) , where

$$b_n = \frac{1}{n^2}, \quad n = 1, 2, \dots$$

To check this, we have to show that

$$\left| \frac{\sin(n^2)}{n^2 + 2^n} \right| \leq \frac{1}{n^2}, \quad \text{for } n = 1, 2, \dots$$

This holds because $|\sin(n^2)| \leq 1$ and

$$n^2 + 2^n \geq n^2, \quad \text{for } n = 1, 2, \dots$$

Since (b_n) is null (by the Power Rule), we deduce, by the Squeeze Rule, that (a_n) is null.

Solution to Exercise D30

(a) (a_n) is a basic null series of type (b), with $c = 0.9$.

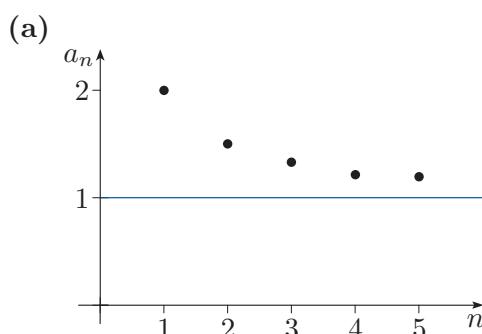
(b) (a_n) is a basic null series of type (d), with $c = 27$.

(c) (a_n) is a basic null series of type (a), with $p = \frac{1}{2}$.

(d) (a_n) is a basic null series of type (e), with $p = 27$.

(e) (a_n) is a basic null series of type (c), with $p = 1$ and $c = \frac{1}{2}$.

Solution to Exercise D31



The sequence diagram suggests that (a_n) converges to 1.

$$(b) b_n = a_n - 1 = \frac{n+1}{n} - 1 = \frac{1}{n}.$$

Hence (b_n) is a null sequence.

Solution to Exercise D32

We have

$$a_n - \frac{1}{2} = \frac{n^3 + 1}{2n^3} - \frac{1}{2} = \frac{1}{2n^3}.$$

We know that $\left(\frac{1}{2n^3}\right)$ is a null sequence, by the

Multiple Rule, so (a_n) converges to $\frac{1}{2}$.

Solution to Exercise D33

In each case we apply Strategy D7.

(a) The dominant term is n^3 , so we write

$$a_n = \frac{n^3 + 2n^2 + 3}{2n^3 + 1} = \frac{1 + 2/n + 3/n^3}{2 + 1/n^3}.$$

Since $(1/n)$ and $(1/n^3)$ are basic null sequences,

$$\begin{aligned} \lim_{n \rightarrow \infty} a_n &= \lim_{n \rightarrow \infty} \frac{1 + 2/n + 3/n^3}{2 + 1/n^3} \\ &= \frac{1 + 0 + 0}{2 + 0} = \frac{1}{2}, \end{aligned}$$

by the Combination Rules.

(b) The dominant term is 3^n , so we write

$$a_n = \frac{n^2 + 2^n}{3^n + n^3} = \frac{n^2/3^n + (2/3)^n}{1 + n^3/3^n}.$$

Since $n^2/3^n = n^2(1/3)^n$ and $n^3/3^n = n^3(1/3)^n$, we see that $(n^2/3^n)$, $((2/3)^n)$ and $(n^3/3^n)$ are all basic null sequences, so

$$\begin{aligned} \lim_{n \rightarrow \infty} a_n &= \lim_{n \rightarrow \infty} \frac{n^2/3^n + (2/3)^n}{1 + n^3/3^n} \\ &= \frac{0 + 0}{1 + 0} = 0, \end{aligned}$$

by the Combination Rules.

(c) The dominant term is $n!$, so we write

$$a_n = \frac{n! + (-1)^n}{2^n + 3n!} = \frac{1 + (-1)^n/n!}{2^n/n! + 3}.$$

Since $((-1)^n/n!)$ and $(2^n/n!)$ are basic null sequences,

$$\begin{aligned} \lim_{n \rightarrow \infty} a_n &= \lim_{n \rightarrow \infty} \frac{1 + (-1)^n/n!}{2^n/n! + 3} \\ &= \frac{1 + 0}{0 + 3} = \frac{1}{3}, \end{aligned}$$

by the Combination Rules.

Solution to Exercise D34

(a) By Rule 5 for rearranging inequalities with $p = n$, we have

$$n^{1/n} \leq 1 + \sqrt{\frac{2}{n-1}} \iff n \leq \left(1 + \sqrt{\frac{2}{n-1}}\right)^n.$$

Using the hint with $x = \sqrt{2/(n-1)}$, we obtain

$$\begin{aligned} \left(1 + \sqrt{\frac{2}{n-1}}\right)^n &\geq \frac{n(n-1)}{2!} \left(\sqrt{\frac{2}{n-1}}\right)^2 \\ &= \frac{n(n-1)}{2} \cdot \frac{2}{n-1} = n. \end{aligned}$$

Thus the right-hand inequality holds for $n \geq 2$, so it follows that the left-hand inequality also holds for $n \geq 2$, as required.

(b) For $n \geq 1$, we have $n^{1/n} \geq 1$. Combining this inequality with that in part (a), we obtain

$$1 \leq n^{1/n} \leq 1 + \sqrt{\frac{2}{n-1}}, \quad \text{for } n \geq 2.$$

Now the sequence (b_n) defined by

$$b_n = \sqrt{\frac{2}{n-1}}, \quad n = 2, 3, \dots$$

is the same as the sequence defined by

$$b_n = \sqrt{\frac{2}{n}}, \quad n = 1, 2, \dots$$

So, since $(1/n)$ is a basic null sequence, it follows from the Multiple Rule and the Power Rule that (b_n) is null. Thus, by the Combination Rules,

$$\lim_{n \rightarrow \infty} \left(1 + \sqrt{\frac{2}{n-1}}\right) = \lim_{n \rightarrow \infty} (1 + b_n) = 1,$$

so, by the Squeeze Rule,

$$\lim_{n \rightarrow \infty} n^{1/n} = 1.$$

Solution to Exercise D35

(a) This sequence is bounded because $1 + (-1)^n$ takes only the values 0 and 2, so

$$|1 + (-1)^n| \leq 2, \quad \text{for } n = 1, 2, \dots$$

(b) This sequence is unbounded. Given any positive number M , there is a positive integer n such that $|(-1)^n n| = n > M$.

(c) This sequence is bounded because

$$\left|\frac{2n+1}{n}\right| = 2 + \frac{1}{n} \leq 3, \quad \text{for } n = 1, 2, \dots$$

Solution to Exercise D36

(a) This sequence is unbounded and hence divergent, by Corollary D15.

(b) This sequence is convergent (with limit 1) and hence bounded, by Theorem D14. In fact,

$$a_n = \frac{n^2 + n}{n^2 + 1} \leq \frac{n^2 + n^2}{n^2} = 2, \quad \text{for } n = 1, 2, \dots$$

(c) This sequence is unbounded and hence divergent, by Corollary D15.

(d) The first few terms of this sequence are

$$1, 2, \frac{1}{3}, 4, \frac{1}{5}, 6, \dots$$

This sequence is unbounded because, given any positive number M , there is an even integer $2n$ such that $(2n)^{(-1)^{2n}} = 2n > M$. Hence the sequence is divergent, by Corollary D15.

Solution to Exercise D37

(a) Each term of (a_n) is positive and

$$\frac{1}{a_n} = \frac{n}{2^n}$$

is a basic null sequence. Hence $a_n \rightarrow \infty$, by the Reciprocal Rule.

(b) The dominant term is 2^n , so we first write

$$2^n - n^9 = 2^n \left(1 - \frac{n^9}{2^n}\right), \quad \text{for } n = 1, 2, \dots$$

Since $(n^9/2^n)$ is a basic null sequence, it follows that $(n^9/2^n)$ is eventually less than 1, so (a_n) is eventually positive.

Next we write

$$\frac{1}{a_n} = \frac{1}{2^n - n^9} = \frac{1/2^n}{1 - n^9/2^n}.$$

Now $(1/2^n)$ and $(n^9/2^n)$ are basic null sequences, so, by the Combination Rules,

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} = \frac{0}{1 - 0} = 0.$$

Hence $a_n \rightarrow \infty$, by the Reciprocal Rule.

(c) We know that $2^n/n \rightarrow \infty$, by part (a), and

$$a_n = 2^n/n + 5n^9 \geq 2^n/n, \quad \text{for } n = 1, 2, \dots$$

Hence $a_n \rightarrow \infty$, by the Squeeze Rule for sequences which tend to infinity.

(Alternatively, you could have used the Reciprocal Rule or the Sum and the Multiple Rules.)

(d) Each term of (a_n) is positive. The dominant term is 2^n , so we write

$$\frac{1}{a_n} = \frac{n^9 + n}{2^n + n^2} = \frac{n^9/2^n + n/2^n}{1 + n^2/2^n}.$$

Now $(n/2^n)$, $(n^2/2^n)$ and $(n^9/2^n)$ are basic null sequences so, by the Combination Rules,

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} = \frac{0+0}{1+0} = 0.$$

Hence $a_n \rightarrow \infty$, by the Reciprocal Rule.

Solution to Exercise D38

(a) (i) $a_2 = 4$, $a_4 = 16$, $a_6 = 36$, $a_8 = 64$, $a_{10} = 100$.

(ii) $a_3 = 9$, $a_7 = 49$, $a_{11} = 121$, $a_{15} = 225$, $a_{19} = 361$.

(iii) $a_1 = 1$, $a_4 = 16$, $a_9 = 81$, $a_{16} = 256$, $a_{25} = 625$.

(b) The first three terms of the odd subsequence are $a_1 = 1$, $a_3 = \frac{1}{3}$, $a_5 = \frac{1}{5}$; the first three terms of the even subsequence are $a_2 = 2$, $a_4 = 4$, $a_6 = 6$.

Solution to Exercise D39

(a) We have

$$a_{2k} = 1 + \frac{1}{2k} \quad \text{and} \quad a_{2k-1} = -1 + \frac{1}{2k-1},$$

for $k = 1, 2, \dots$. Thus

$$\lim_{k \rightarrow \infty} a_{2k} = 1, \quad \text{whereas} \quad \lim_{k \rightarrow \infty} a_{2k-1} = -1.$$

Hence (a_n) is divergent, by the First Subsequence Rule.

(b) We have

$$a_{3k} = k - \lfloor k \rfloor = 0$$

and

$$a_{3k+1} = k + \frac{1}{3} - \lfloor k + \frac{1}{3} \rfloor = k + \frac{1}{3} - k = \frac{1}{3},$$

for $k = 1, 2, \dots$. Thus

$$\lim_{k \rightarrow \infty} a_{3k} = 0, \quad \text{whereas} \quad \lim_{k \rightarrow \infty} a_{3k+1} = \frac{1}{3}.$$

Hence (a_n) is divergent, by the First Subsequence Rule.

(c) We have

$$a_1 = 1, a_2 = 0, a_3 = -3,$$

$$a_4 = 0, a_5 = 5, a_6 = 0, \dots$$

Now

$$\begin{aligned} a_{4k+1} &= (4k+1) \sin(2k\pi + \frac{1}{2}\pi) \\ &= 4k+1, \end{aligned}$$

for $k = 1, 2, \dots$. Thus $a_{4k+1} \rightarrow \infty$ as $k \rightarrow \infty$.

Hence (a_n) is divergent, by the Second Subsequence Rule.

Solution to Exercise D40

Since (a_n) is increasing, it follows from the Monotonic Sequence Theorem that either (a_n) is convergent or $a_n \rightarrow \infty$.

If $a_n \rightarrow \infty$ as $n \rightarrow \infty$, then it follows from Theorem D19(b) that $a_{n_k} \rightarrow \infty$ as $k \rightarrow \infty$, and we know that this is false.

Hence (a_n) must be convergent.

Unit D3

Series

Introduction

In this unit you will study *infinite series*. Informally, an infinite series is the sum of infinitely many numbers such as

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots$$

Although in everyday language the words ‘sequence’ and ‘series’ are often used interchangeably, in mathematics they represent distinct, but related, concepts.

As with sequences, an infinite series may converge or diverge. The unit begins with a formal definition of a convergent infinite series, illustrated with several examples. You will then learn how to use various tests to determine whether or not a series converges.

Because the study of series involves studying related sequences, this unit depends heavily on the ideas and results of Unit D2 *Sequences*; therefore, before studying this unit you should make sure that you understand Unit D2, Sections 1, 2 and 3, and that you are familiar with Sections 4 and 5.

When working with series, we often need to determine whether a related sequence converges and, if it does, find the value of its limit. In this unit we do not always give as much detail in such calculations as we did in Unit D2. As is the case throughout this module, the amount of detail given in solutions to the exercises and worked exercises indicates the amount you should give in your own solutions.

1 Introducing series

In this section you will see how a convergent infinite series can be defined formally in terms of convergent sequences. You will meet several examples of such series and discover various properties that all convergent series have in common.

1.1 What is a convergent series?

We begin by considering a paradox of Zeno.

The Ancient Greek philosopher and mathematician Zeno lived and worked in Elea in southern Italy during the 5th century BCE. Zeno proposed a number of paradoxes of the infinite, which have intrigued succeeding generations. His paradoxes include ‘The Flying Arrow’ and ‘Achilles and the Tortoise’.

In one of his paradoxes, Zeno claimed that it is impossible for an object to travel a given distance, since it must first travel half the distance, then half of the remaining distance, then half of what remains, and so on. There must always remain some distance left to travel, so the journey cannot be completed.

This paradox relies partly on the intuitive feeling that it is impossible to add up an infinite number of positive quantities and obtain a finite answer. However, the illustration of the paradox in Figure 1 suggests that, in this case, a finite answer is certainly plausible.

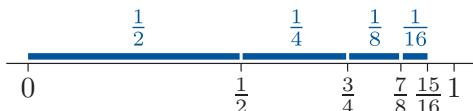


Figure 1 The distances in Zeno's paradox

The distance from 0 to 1 can be split up into the infinite sequence of distances $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$, each distance being half of the preceding one, so it seems reasonable to write

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots = 1.$$

We now give a justification for this statement. Let s_n be the sum of the first n terms on the left-hand side. We call this the *nth partial sum*. Then

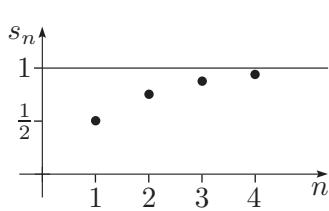


Figure 2 The sequence (s_n) of partial sums

The sequence diagram for the sequence (s_n) of partial sums is shown in Figure 2.

To obtain the *nth* partial sum, we use the formula for the sum of a finite geometric series. Here is a reminder of the formula.

Sum of a finite geometric series

The geometric series with first term a , common ratio $r \neq 1$ and n terms has the sum

$$a + ar + \dots + ar^{n-1} = \frac{a(1 - r^n)}{1 - r}.$$

By applying this formula in the case that $a = r = \frac{1}{2}$, we obtain

$$\begin{aligned} s_n &= \frac{1}{2} + \frac{1}{4} + \dots + \left(\frac{1}{2}\right)^n \\ &= \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2} \left(\frac{1}{2}\right)^{n-1} \\ &= \frac{\frac{1}{2} \left(1 - \left(\frac{1}{2}\right)^n\right)}{1 - \frac{1}{2}} = 1 - \left(\frac{1}{2}\right)^n. \end{aligned}$$

The sequence $\left(\left(\frac{1}{2}\right)^n\right)$ is a basic null sequence, so

$$\lim_{n \rightarrow \infty} s_n = 1 - \lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n = 1.$$

It is this precise mathematical statement that $s_n \rightarrow 1$ as $n \rightarrow \infty$ which justifies our informal statement earlier that

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots = 1.$$

We now use this approach to define a *convergent infinite series*.

Definitions

Let (a_n) be a sequence. Then the expression

$$a_1 + a_2 + a_3 + \dots$$

is an **infinite series**, or simply a **series**.

We call a_n the **n th term** of the series and

$$s_n = a_1 + a_2 + \dots + a_n$$

the **n th partial sum** of the series.

The series is **convergent** with **sum** s (or **converges to** s) if its sequence (s_n) of partial sums converges to s . In this case, we write

$$a_1 + a_2 + a_3 + \dots = s.$$

The series **diverges**, or is **divergent**, if the sequence (s_n) diverges.

Notice that the sum of a convergent infinite series is the limit of the sequence of its partial sums. Thus we can prove results about a series by applying known results about sequences to its partial sums (s_n) .

Remarks

- When you have a series $a_1 + a_2 + \dots$, it is important to distinguish between the sequence of terms

$$(a_n) = a_1, a_2, a_3, \dots$$

and the sequence of partial sums

$$(s_n) = a_1, a_1 + a_2, a_1 + a_2 + a_3, \dots$$

It may help you remember the difference if you think of s_n as the *sum* of the series up to and including the n th term (s standing for ‘sum’), and a_n as the amount that is *added* to the series by the n th term (a standing for ‘added’).

It may also be helpful to think of the series as a ‘snake’ which

- has length s_n on its n th birthday, and
- grows by the length a_n in its n th year.

This is illustrated in Figure 3.

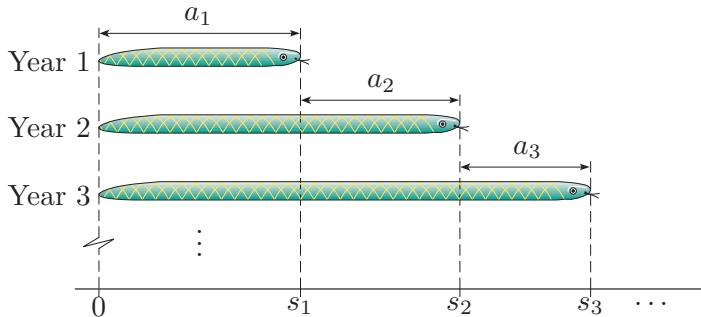


Figure 3 The sequences (s_n) and (a_n)

This picture has its limitations, however; for example, it assumes that the snake lives forever. Also the terms a_n need not be positive, so the snake may shrink or even have negative length!

2. Sometimes the first term in a series may not be a_1 ; for example, it might instead be a_0 or a_3 . However, when a series begins with a term other than a_1 , we still calculate the n th partial sum s_n by adding all the terms up to and including a_n .

For example, if the first term of a series is a_0 , then the n th partial sum is

$$s_n = a_0 + a_1 + \cdots + a_n.$$

Hence, in the case of a series starting with a_0 , there is a 0th partial sum $s_0 = a_0$, and the n th partial sum is the sum of $n + 1$ terms.

3. Note that changing, deleting or adding a *finite* number of terms does not affect the convergence of a series, but may affect its sum. For example, the series

$$1 + 2 + 3 + 4 + 5 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$$

is convergent with sum $1 + 2 + 3 + 4 + 5 + 1 = 16$.

Worked Exercise D28

For each of the following series, calculate the n th partial sum and determine whether the series is convergent or divergent.

- (a) $1 + 1 + 1 + \cdots$ (b) $\frac{1}{3} + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^3 + \cdots$ (c) $2 + 4 + 8 + \cdots$

Solution

- (a) In this case,

$$s_n = \underbrace{1 + 1 + \cdots + 1}_{n \text{ terms}} = n.$$

So $s_n \rightarrow \infty$ as $n \rightarrow \infty$, and this series is divergent.

- (b) Here we can use the formula for the sum of a finite geometric series to obtain the n th partial sum.

Putting $a = r = \frac{1}{3}$ in the formula for the sum of a finite geometric series, we obtain

$$\begin{aligned}s_n &= \frac{1}{3} + \left(\frac{1}{3}\right)^2 + \cdots + \left(\frac{1}{3}\right)^n \\ &= \frac{\frac{1}{3}(1 - (\frac{1}{3})^n)}{1 - \frac{1}{3}} = \frac{1}{2}(1 - (\frac{1}{3})^n).\end{aligned}$$

Since $((\frac{1}{3})^n)$ is a basic null sequence,

$$\lim_{n \rightarrow \infty} s_n = \frac{1}{2},$$

so this series is convergent with sum $\frac{1}{2}$.

- (c) In this case, the formula for the sum of a finite geometric series with $a = r = 2$ gives

$$s_n = 2 + 4 + 8 + \cdots + 2^n = \frac{2(1 - 2^n)}{1 - 2} = 2^{n+1} - 2.$$

So $s_n \rightarrow \infty$ as $n \rightarrow \infty$, and this series is divergent.

The next worked exercise shows that infinite series need to be manipulated with care.

Worked Exercise D29

Explain why the following proof is incorrect.

Claim (incorrect!)

$$2 + 4 + 8 + \cdots = -2.$$

Proof (incorrect!) Let

$$2 + 4 + 8 + \cdots = s.$$

Multiplying through by $\frac{1}{2}$ gives

$$1 + 2 + 4 + 8 + \cdots = \frac{1}{2}s,$$

which can be written as

$$1 + s = \frac{1}{2}s.$$

So $\frac{1}{2}s = -1$ and hence

$$s = -2.$$



Solution

The problem with this ‘proof’ is that it is based on the assumption that s is a finite number. In fact, as we saw in Worked Exercise D28(c), the series

$$2 + 4 + 8 + \dots$$

is divergent and its partial sums tend to infinity. Thus we cannot perform arithmetic operations with s and so the proof is not valid.

We can avoid reaching absurd conclusions such as that in Worked Exercise D29 by performing arithmetic operations only with infinite series which we know to be *convergent*. It is therefore very important to be able to check whether or not a given series is convergent. You will meet many ways of doing this for different types of series in this unit.

Sigma notation

First, however, we show how to use *sigma notation* as a convenient way to represent infinite series.

From your previous studies you will be familiar with the use of sigma notation as a shorthand way of writing *finite* sums. For example, instead of $a_1 + a_2 + \dots + a_{10}$, we can write

$$\sum_{n=1}^{10} a_n,$$

where the symbol \sum is the Greek upper-case letter sigma (standing for ‘sum’) and the subscript n takes all integer values from 1 to 10 inclusive.

This notation can readily be adapted to represent infinite series. Instead of $a_1 + a_2 + a_3 + \dots$, we write

$$\sum_{n=1}^{\infty} a_n,$$

which is read as ‘sigma, $n = 1$ to infinity, a_n ’, or ‘the sum from $n = 1$ to infinity of a_n ’. For example,

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots$$

Remarks

- An alternative layout for sigma notation is $\sum_{n=1}^{\infty} a_n$. We sometimes use the simpler notation $\sum a_n$ to denote a general infinite series with terms a_n .
- Note that there is no term a_{∞} in the series $\sum_{n=1}^{\infty} a_n$. The symbol ∞ is here used to mean that the subscript n takes every integer value greater than or equal to 1.
- When using sigma notation to represent the n th partial sum s_n of a series, we use another letter for the subscript of the terms to avoid n having two different meanings in the same expression; we may write

$$s_n = a_1 + a_2 + \cdots + a_n = \sum_{k=1}^n a_k.$$

For example, we write

$$s_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = \sum_{k=1}^n \frac{1}{k}.$$

The letters i, j, k, l, m, n, p and q are commonly used for subscript variables.

- If a series begins with a term other than a_1 , then we adapt the notation appropriately; for instance,

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \cdots \quad \text{or} \quad \sum_{n=3}^{\infty} a_n = a_3 + a_4 + a_5 + \cdots.$$

For example, we write

$$\sum_{n=0}^{\infty} r^n = 1 + r + r^2 + \cdots.$$

Exercise D41

For each of the following series, calculate the n th partial sum s_n and determine whether the series is convergent or divergent. If it is convergent, find its sum.

$$(a) \sum_{n=1}^{\infty} \left(-\frac{1}{3}\right)^n \quad (b) \sum_{n=1}^{\infty} (-1)^n \quad (c) \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n$$

The series considered so far in this section are all geometric series. In general, the **(infinite) geometric series** with first term a and common ratio r is

$$\sum_{n=0}^{\infty} ar^n = a + ar + ar^2 + \dots .$$

Note that it is conventional to regard the term ar^n as the n th term of a geometric series, which means that the summation goes from $n = 0$ to ∞ . (We did not adopt this convention for some of the geometric series you have met so far, but we will do so from now on.)

The following theorem enables us to decide whether any given geometric series is convergent or divergent.

Theorem D24 Geometric series

(a) If $|r| < 1$, then $\sum_{n=0}^{\infty} ar^n$ is convergent, with sum $\frac{a}{1-r}$.

(b) If $|r| \geq 1$ and $a \neq 0$, then $\sum_{n=0}^{\infty} ar^n$ is divergent.

Proof (a) If $r \neq 1$, then the n th partial sum s_n is given by the formula for the sum of a finite geometric series with $n+1$ terms, so

$$s_n = a + ar + ar^2 + \dots + ar^n = \frac{a(1 - r^{n+1})}{1 - r}. \quad (1)$$

Now, if $|r| < 1$, then (r^n) is a basic null sequence, so

$$\begin{aligned} \lim_{n \rightarrow \infty} s_n &= \lim_{n \rightarrow \infty} \frac{a(1 - r^{n+1})}{1 - r} \\ &= \frac{a}{1 - r} \left(1 - \lim_{n \rightarrow \infty} r^{n+1}\right) = \frac{a}{1 - r}, \end{aligned}$$

by the Combination Rules for sequences that you met in Unit D2. Thus, if $|r| < 1$, then

$$\sum_{n=0}^{\infty} ar^n \text{ is convergent, with sum } \frac{a}{1-r}.$$

(b) We deal separately with the cases $r = \pm 1$ and $|r| > 1$.

We saw in Worked Exercise D28(a) and Exercise D41(b) that a geometric series with $r = \pm 1$ and $a \neq 0$ is divergent. The same arguments can be used to show that any geometric series with $r = \pm 1$ is divergent, whatever the value of a .

If $|r| > 1$, then (s_n) is given by equation (1) and $|r|^{n+1} \rightarrow \infty$ as $n \rightarrow \infty$, so the sequence (s_n) is unbounded and hence divergent.

Thus, if $|r| \geq 1$, then $\sum_{n=0}^{\infty} ar^n$ is divergent. ■

1.2 Telescoping series

Geometric series are easy to deal with because there is a formula for the n th partial sum s_n . The next exercise concerns another series for which we can calculate a formula for s_n .

Exercise D42

Calculate the first four partial sums of the following series, giving your answers as fractions:

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \frac{1}{3 \times 4} + \dots$$

The partial sums obtained in Exercise D42 suggest the general formula

$$s_n = \frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \dots + \frac{1}{n(n+1)} = \frac{n}{n+1}.$$

This formula can be proved by using the identity

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}, \quad \text{for } n = 1, 2, \dots,$$

which implies that

$$s_n = \sum_{k=1}^n \frac{1}{k(k+1)} = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right).$$

Thus

$$\begin{aligned} s_n &= \frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \dots + \frac{1}{(n-1)n} + \frac{1}{n(n+1)} \\ &= \left(\frac{1}{1} - \frac{1}{2} \right) + \left(\frac{1}{2} - \frac{1}{3} \right) + \dots + \left(\frac{1}{n-1} - \frac{1}{n} \right) + \left(\frac{1}{n} - \frac{1}{n+1} \right) \\ &= 1 - \frac{1}{n+1} = \frac{n}{n+1}. \end{aligned}$$

(This series is said to be *telescoping* because of the cancellation of the adjacent terms $-\frac{1}{2}, \frac{1}{2}, -\frac{1}{3}, \frac{1}{3}, \dots$)

Since

$$\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \frac{n}{n+1} = \lim_{n \rightarrow \infty} \frac{1}{1 + 1/n} = 1,$$

we deduce that the given series is convergent, with sum

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1.$$

Exercise D43

Find the n th partial sum s_n of $\sum_{n=1}^{\infty} \frac{1}{n(n+2)}$, using the identity

$$\frac{2}{n(n+2)} = \frac{1}{n} - \frac{1}{n+2}, \quad \text{for } n = 1, 2, \dots$$

Deduce that this series is convergent and find its sum.

1.3 Combination Rules for series

You have already seen that performing arithmetic operations on the divergent series $2 + 4 + 8 + \dots$ can lead to absurd conclusions. However, we can perform arithmetic operations on convergent series. The following result shows that there are Combination Rules for convergent series, which follow directly from the Combination Rules for sequences.

Theorem D25 Combination Rules for convergent series

Suppose that $\sum_{n=1}^{\infty} a_n = s$ and $\sum_{n=1}^{\infty} b_n = t$. Then

Sum Rule $\sum_{n=1}^{\infty} (a_n + b_n) = s + t$

Multiple Rule $\sum_{n=1}^{\infty} \lambda a_n = \lambda s$, for any real number λ .

Proof Consider the sequences of partial sums (s_n) and (t_n) , where

$$s_n = \sum_{k=1}^n a_k \quad \text{and} \quad t_n = \sum_{k=1}^n b_k.$$

We know that $s_n \rightarrow s$ as $n \rightarrow \infty$ and $t_n \rightarrow t$ as $n \rightarrow \infty$.

Sum Rule The n th partial sum of the series $\sum_{n=1}^{\infty}(a_n + b_n)$ is

$$\begin{aligned}\sum_{k=1}^n(a_k + b_k) &= (a_1 + b_1) + (a_2 + b_2) + \cdots + (a_n + b_n) \\ &= (a_1 + a_2 + \cdots + a_n) + (b_1 + b_2 + \cdots + b_n) \\ &= s_n + t_n.\end{aligned}$$

By the Sum Rule for sequences,

$$\lim_{n \rightarrow \infty}(s_n + t_n) = \lim_{n \rightarrow \infty}s_n + \lim_{n \rightarrow \infty}t_n = s + t,$$

so the sequence $(s_n + t_n)$ of partial sums of $\sum_{n=1}^{\infty}(a_n + b_n)$ has limit $s + t$.

Hence this series is convergent and

$$\sum_{n=1}^{\infty}(a_n + b_n) = s + t.$$

Multiple Rule The n th partial sum of the series $\sum_{n=1}^{\infty}\lambda a_n$ is

$$\begin{aligned}\sum_{k=1}^n\lambda a_k &= \lambda a_1 + \lambda a_2 + \cdots + \lambda a_n \\ &= \lambda(a_1 + a_2 + \cdots + a_n) \\ &= \lambda s_n.\end{aligned}$$

By the Multiple Rule for sequences,

$$\lim_{n \rightarrow \infty}(\lambda s_n) = \lambda \lim_{n \rightarrow \infty}s_n = \lambda s,$$

so the sequence (λs_n) of partial sums of $\sum_{n=1}^{\infty}\lambda a_n$ has limit λs . Hence this series is convergent and

$$\sum_{n=1}^{\infty}\lambda a_n = \lambda s.$$



Worked Exercise D30

Prove that the following series is convergent and calculate its sum:

$$\sum_{n=1}^{\infty}\left(\frac{1}{2^n} + \frac{3}{n(n+1)}\right).$$

Solution

💡 This series is of the form $\sum_{n=1}^{\infty}(a_n + 3b_n)$, where

$$a_n = \frac{1}{2^n} \quad \text{and} \quad b_n = \frac{1}{n(n+1)}.$$

We studied the series $\sum_{n=1}^{\infty}a_n$ and $\sum_{n=1}^{\infty}b_n$ earlier in the section. 💡

At the beginning of Subsection 1.1 we saw that

$$\sum_{n=1}^{\infty} \frac{1}{2^n} \text{ is convergent, with sum 1.}$$

Also, in Subsection 1.2, we saw that

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} \text{ is convergent, with sum 1.}$$

To deduce the sum of the given series from these results, we use both the Sum Rule and the Multiple Rule.

Hence, by the Combination Rules,

$$\sum_{n=1}^{\infty} \left(\frac{1}{2^n} + \frac{3}{n(n+1)} \right) \text{ is convergent, with sum } 1 + (3 \times 1) = 4.$$

Exercise D44

Prove that the following series is convergent and calculate its sum:

$$\sum_{n=1}^{\infty} \left(\left(-\frac{3}{4} \right)^n - \frac{2}{n(n+1)} \right).$$

We conclude this subsection with the following corollary to the Multiple Rule for convergent series, which tells us that a non-zero multiple of a *divergent* series is also divergent. (Remember, though, that you can *never* perform arithmetic operations with divergent series.)

Corollary D26 Multiple Rule for divergent series

Suppose that $\sum_{n=1}^{\infty} a_n$ is divergent and that λ is a non-zero real number.

Then $\sum_{n=1}^{\infty} \lambda a_n$ is divergent.

Proof We use proof by contradiction. Suppose that $\sum_{n=1}^{\infty} a_n$ is divergent, and that λ is some non-zero real number such that $\sum_{n=1}^{\infty} \lambda a_n$ is convergent.

Then it follows from the Multiple Rule for convergent series with multiple $1/\lambda$ that

$$\sum_{n=1}^{\infty} a_n = \frac{1}{\lambda} \sum_{n=1}^{\infty} \lambda a_n \text{ is convergent.}$$

However, this is a contradiction since we know that $\sum_{n=1}^{\infty} a_n$ is divergent. So our original assumption was wrong and we must in fact have

$$\sum_{n=1}^{\infty} \lambda a_n \text{ is divergent, for any non-zero real number } \lambda.$$



1.4 Non-null Test

For all the infinite series we have so far considered, it is possible to derive a simple formula for the n th partial sum. For many series, however, this is difficult or even impossible.

Nevertheless, it may still be possible to decide whether such series are convergent or divergent by applying various tests. Our first test arises from the following result.

Theorem D27

If $\sum_{n=1}^{\infty} a_n$ is a convergent series, then its sequence of terms (a_n) is a null sequence.

Proof Let

$$s_n = \sum_{k=1}^n a_k = a_1 + a_2 + \cdots + a_n$$

denote the n th partial sum of the series $\sum_{n=1}^{\infty} a_n$. Because $\sum_{n=1}^{\infty} a_n$ is convergent, we know that (s_n) is a convergent sequence, with limit s , say.

We want to deduce that (a_n) is null. To do this, we note that

$$a_n = s_n - s_{n-1}, \quad \text{for } n \geq 2.$$

Thus, by the Combination Rules for convergent sequences,

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} (s_n - s_{n-1}) = \lim_{n \rightarrow \infty} s_n - \lim_{n \rightarrow \infty} s_{n-1}.$$

The sequence $(s_{n-1})_2^{\infty}$ is the same as the sequence $(s_n)_1^{\infty}$, so $s_{n-1} \rightarrow s$ as $n \rightarrow \infty$. Thus

$$\lim_{n \rightarrow \infty} a_n = s - s = 0$$

so (a_n) is a null sequence, as required.



The following test for divergence is an immediate corollary of Theorem D27.

Corollary D28 Non-null Test

If (a_n) is not a null sequence, then $\sum_{n=1}^{\infty} a_n$ is divergent.

Although it is sometimes obvious that a sequence (a_n) is not null, it can be useful to have a method for showing this. You saw in Unit D2 that if (a_n) is a null sequence, then $(|a_n|)$ is also null, as are all subsequences of $(|a_n|)$. This leads to the following strategy.

Strategy D9

To show that $\sum_{n=1}^{\infty} a_n$ is divergent using the Non-null Test, check that the sequence (a_n) is not null by showing that $(|a_n|)$ has either

- (a) a convergent subsequence with non-zero limit,
- or
- (b) a subsequence which tends to infinity.

Often when we apply Strategy D9 there is no need to consider a subsequence, because the *whole* sequence $(|a_n|)$ tends to a non-zero limit or to infinity, as in the following Worked Exercise.

Worked Exercise D31

Prove that each of the following series is divergent.

$$(a) \sum_{n=1}^{\infty} (-1)^n \quad (b) \sum_{n=1}^{\infty} \frac{2^n}{n}$$

Solution

(a) Let $a_n = (-1)^n$. Then $|a_n| = 1$.

 We apply Strategy D9 to the whole sequence $|a_n|$. 

So, since

$$\lim_{n \rightarrow \infty} |a_n| = 1 \neq 0,$$

the sequence (a_n) is not null. Hence, by the Non-null Test,

$$\sum_{n=1}^{\infty} (-1)^n \text{ is divergent.}$$

(b) Let $a_n = 2^n/n$.

Here there is no need to consider $|a_n|$ as a_n is positive. The dominant term in the formula for a_n is in the numerator, so we use the Reciprocal Rule from Unit D2 to show that $a_n \rightarrow \infty$.

Then $(1/a_n) = (n/2^n)$ is a basic null sequence. So, by the Reciprocal Rule for sequences, $a_n \rightarrow \infty$ as $n \rightarrow \infty$. Hence, by the Non-null Test,

$$\sum_{n=1}^{\infty} \frac{2^n}{n} \text{ is divergent.}$$

Exercise D45

Prove that $\sum_{n=1}^{\infty} \frac{(-1)^{n+1} n^2}{2n^2 + 1}$ is divergent.

Note that the converse of the Non-null Test is *false*. If the sequence (a_n) is null, then it is not necessarily true that the series $\sum_{n=1}^{\infty} a_n$ is convergent. For example, the sequence $(1/n)$ is null, but the series

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots \text{ is divergent.}$$

(We prove this rather surprising fact at the beginning of the next section.) So you can *never* use the Non-null Test to prove that a series is convergent.

2 Series with non-negative terms

In this section we restrict our attention to series $\sum_{n=1}^{\infty} a_n$ with *non-negative* terms. In other words, we assume that

$$a_n \geq 0, \quad \text{for } n = 1, 2, \dots$$

It follows that the partial sums of $\sum_{n=1}^{\infty} a_n$, given by

$$s_1 = a_1$$

$$s_2 = a_1 + a_2$$

$$\vdots$$

$$s_n = a_1 + a_2 + \dots + a_n,$$

form an *increasing* sequence. This is illustrated in Figure 4 and in the sequence diagram in Figure 5.

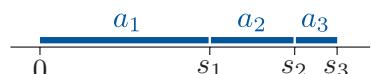


Figure 4 A series with non-negative terms

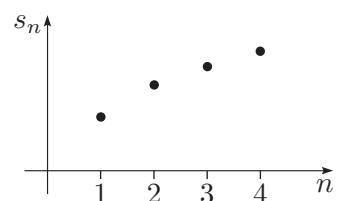


Figure 5 The partial sums of a series with non-negative terms

As in Section 1, we are interested in finding out whether a series with non-negative terms is convergent, even if we are unable to evaluate its sum or partial sums. The fact that the sequence (s_n) of partial sums is *increasing* helps us to determine whether (s_n) is convergent, since we can use the Monotone Convergence Theorem which you met in Unit D2. We restate this result below.

Theorem D29 Monotone Convergence Theorem

If the sequence (a_n) is either

- increasing and bounded above, or
- decreasing and bounded below,

then (a_n) is convergent.

Thus, if we can prove that the sequence (s_n) of partial sums of a series with non-negative terms is bounded above, then it follows from the Monotone Convergence Theorem that (s_n) is convergent, and hence that the series $\sum_{n=1}^{\infty} a_n$ is convergent.

In this section you will meet several tests that can be used to check whether a series with non-negative terms is convergent and also several examples of such series.

2.1 Tests for convergence

We begin by studying two important examples of series with non-negative terms:

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots$$

Exercise D46

Use a calculator to find the first eight partial sums of each of the series $\sum_{n=1}^{\infty} \frac{1}{n}$ and $\sum_{n=1}^{\infty} \frac{1}{n^2}$ (giving your answers to two decimal places), and plot your answers on a sequence diagram.

From the solution to Exercise D46, it appears that

$$\sum_{n=1}^{\infty} \frac{1}{n} \text{ diverges} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^2} \text{ converges.}$$

We now show that this is indeed the case. To do this, we look carefully at the partial sums of these two series to see whether the sequences of their partial sums converge or diverge.

We begin by considering $\sum_{n=1}^{\infty} \frac{1}{n}$. This series is often called the *harmonic* series, since its terms are proportional to the lengths of strings that produce harmonic tones in music.

The earliest recorded proof that the harmonic series is divergent is in a treatise dating from c.1350 by the French medieval philosopher Nicole Oresme (c.1320–1382).

Worked Exercise D32

Prove that the series $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent.

Solution

 Notice that the sequence $(1/n)$ is null, so we cannot use the Non-null Test here. The key step in the proof is to find a subsequence of partial sums that tends to infinity. We do this by arranging the terms of the partial sums into groups. 

Let s_n be the n th partial sum of the series. Then:

$$\begin{aligned}s_n &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10} + \cdots + \frac{1}{n} \\&= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4} \right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \right) + \cdots \\&\quad \cdots + \left(\frac{1}{2^k+1} + \frac{1}{2^k+2} + \cdots + \frac{1}{2^k+2^k} \right) + \cdots.\end{aligned}$$

 The k th bracket contains 2^k terms, each at least equal to $\frac{1}{2^k+2^k}$, so the sum of the terms in each bracket is at least equal to

$\frac{2^k}{2^k+2^k} = \frac{1}{2}$. This fact enables us to use the Squeeze Rule for sequences which tend to infinity, which you met in Unit D2. 

It follows that the subsequence (s_{2^k}) of partial sums is:

$$\begin{aligned}s_1 &= 1, \\s_2 &= 1 + \frac{1}{2}, \\s_4 &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4} \right) > 1 + \frac{1}{2} + \frac{1}{2}, \\s_8 &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4} \right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \right) > 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}, \\&\vdots \\s_{2^k} &> 1 + \underbrace{\frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2}}_{k \text{ terms}} = 1 + \frac{1}{2}k.\end{aligned}$$

Since $1 + \frac{1}{2}k \rightarrow \infty$ as $k \rightarrow \infty$, it follows from the Squeeze Rule that $s_{2^k} \rightarrow \infty$.

 We now use the Second Subsequence Rule from Unit D2, which says that if a sequence has a subsequence that tends to infinity, then the sequence is divergent. 

Hence the sequence (s_n) is divergent by the Second Subsequence Rule. It follows that the series $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent.

We now consider the second series from Exercise D46.

Worked Exercise D33

Prove that the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent.

Solution

 All the terms of the series are positive, so the sequence of partial sums is increasing. We show that the sequence of partial sums is bounded above and use the Monotone Convergence Theorem. 

Let s_n be the n th partial sum of the series. Then, using the method of telescoping cancellation and the fact that

$$\frac{1}{k^2} < \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$$

for all integers $k > 1$, we have

$$\begin{aligned} s_n &= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots + \frac{1}{n^2} \\ &< 1 + \frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \frac{1}{3 \times 4} + \cdots + \frac{1}{(n-1) \times n} \\ &= 1 + \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 2 - \frac{1}{n} \\ &< 2. \end{aligned}$$

It follows that (s_n) is both increasing and bounded above (by 2, for example) so that, by the Monotone Convergence Theorem, (s_n) is convergent. So the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent.

In fact, remarkably, it can be shown that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

However, a proof of this (which depends on use of trigonometric series or of complex analysis) goes beyond the scope of this module.

The problem of finding the exact sum of the series $\sum_{n=1}^{\infty} (1/n^2)$ is known as the Basel problem. The problem was first posed by the Italian mathematician Pietro Mengoli (1626–1686) in 1644. After withstanding the attack of many leading mathematicians, including several members of the Bernoulli family, it was solved by the young Leonhard Euler (1707–1783) in 1734, bringing him immediate fame. The problem is named after the city of Basel in Switzerland, the hometown of Euler and of the Bernoulli family.

Just as with sequences, there are some general results that enable us to deduce the convergence or divergence of a given series from the known convergence or divergence of another series. Our next result is of this type.

Theorem D30 Comparison Test

(a) If

$$0 \leq a_n \leq b_n, \quad \text{for } n = 1, 2, \dots, \tag{2}$$

and $\sum_{n=1}^{\infty} b_n$ is convergent, then $\sum_{n=1}^{\infty} a_n$ is convergent.

(b) If

$$0 \leq b_n \leq a_n, \quad \text{for } n = 1, 2, \dots, \tag{3}$$

and $\sum_{n=1}^{\infty} b_n$ is divergent, then $\sum_{n=1}^{\infty} a_n$ is divergent.

Proof (a) Assume that inequality (2) holds. Then the n th partial sums

$$s_n = a_1 + a_2 + \cdots + a_n, \quad n = 1, 2, \dots,$$

and

$$t_n = b_1 + b_2 + \cdots + b_n, \quad n = 1, 2, \dots,$$

satisfy

$$s_n \leq t_n, \quad \text{for } n = 1, 2, \dots.$$

We also know that $\sum_{n=1}^{\infty} b_n$ is convergent, so the increasing sequence (t_n) is convergent with limit t , say. Hence

$$s_n \leq t_n \leq t, \quad \text{for } n = 1, 2, \dots,$$

so the increasing sequence (s_n) is bounded above by t . By the Monotone Convergence Theorem, (s_n) is also convergent, so $\sum_{n=1}^{\infty} a_n$ is a convergent series.

- (b) Now assume that inequality (3) holds. Notice that the statement

$$\text{if } \sum_{n=1}^{\infty} b_n \text{ is divergent, then } \sum_{n=1}^{\infty} a_n \text{ is divergent}$$

is equivalent to the statement

$$\text{if } \sum_{n=1}^{\infty} a_n \text{ is convergent, then } \sum_{n=1}^{\infty} b_n \text{ is convergent.}$$

 The second statement is the *contrapositive* of the first statement and so is equivalent to it, as you saw in Unit A3 *Mathematical language and proof*. 

But the second statement follows from part (a) with the roles of a_n and b_n interchanged, so this proves part (b). 

Remarks

1. Informally, in part (a) we say that $\sum_{n=1}^{\infty} a_n$ ‘is dominated by’ $\sum_{n=1}^{\infty} b_n$; and in part (b) that $\sum_{n=1}^{\infty} a_n$ ‘dominates’ $\sum_{n=1}^{\infty} b_n$.
2. In the proof of part (a), if we apply the Limit Inequality Rule from Unit D2 to the inequality

$$s_n \leq t_n, \quad \text{for } n = 1, 2, \dots,$$

then we can in fact deduce that $\sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} b_n$.

3. When using the Comparison Test, it is sufficient to show that the necessary inequalities in parts (a) and (b) hold *eventually*; that is, that $0 \leq a_n \leq b_n$ or $0 \leq b_n \leq a_n$ for all $n > N$, for some number N . (You met this idea of a property holding *eventually* in Unit D2.)

In applications, we use the Comparison Test in the following way – sometimes informally called a ‘guess then check’ approach.

Strategy D10

- To show that a series $\sum_{n=1}^{\infty} a_n$ of non-negative terms is *convergent* using the Comparison Test, do the following.
 1. Guess that $\sum_{n=1}^{\infty} a_n$ is dominated by a convergent series $\sum_{n=1}^{\infty} b_n$.
 2. Check that $0 \leq a_n \leq b_n$, for $n = 1, 2, \dots$
- To show that a series $\sum_{n=1}^{\infty} a_n$ of non-negative terms is *divergent* using the Comparison Test, do the following.
 1. Guess that $\sum_{n=1}^{\infty} a_n$ dominates a divergent series $\sum_{n=1}^{\infty} b_n$.
 2. Check that $0 \leq b_n \leq a_n$, for $n = 1, 2, \dots$

In either case, the first step is to find a suitable series $\sum_{n=1}^{\infty} b_n$ to compare with the series $\sum_{n=1}^{\infty} a_n$ that we are investigating. To do this, we choose a series whose n th term b_n seems likely to be greater than or equal to a_n (less than or equal to a_n) and which we know converges (diverges).

Carrying out the check in step 2 of Strategy D10 will then show whether or not our guess of $\sum_{n=1}^{\infty} b_n$ was suitable. This is illustrated in the next two worked exercises.

Worked Exercise D34

Prove that the series $\sum_{n=1}^{\infty} \frac{1}{n^3}$ is convergent.

Solution

 We need to show that $\sum_{n=1}^{\infty} \frac{1}{n^3}$ is dominated by a series which we know to be convergent. 

We guess that $\sum_{n=1}^{\infty} \frac{1}{n^3}$ is dominated by $\sum_{n=1}^{\infty} \frac{1}{n^2}$. Indeed, since

$$n^3 \geq n^2 \geq 0, \quad \text{for } n = 1, 2, \dots,$$

we have

$$0 \leq \frac{1}{n^3} \leq \frac{1}{n^2}, \quad \text{for } n = 1, 2, \dots$$

Since $\sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent (by Worked Exercise D33), we deduce from

part (a) of the Comparison Test that $\sum_{n=1}^{\infty} \frac{1}{n^3}$ is convergent.

There is no simple way of writing the sum of the series $\sum_{n=1}^{\infty} \frac{1}{n^3}$. This value is now known as Apéry's constant after the French mathematician Roger Apéry who in 1978 proved that it is irrational. Apéry's constant arises naturally in a number of physical problems; for example, in the quantification of black body energy radiation.

Worked Exercise D35

Prove that the series $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ is divergent.

Solution

We guess that $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ dominates $\sum_{n=1}^{\infty} \frac{1}{n}$. Indeed, since

$$n \geq \sqrt{n} \geq 0, \quad \text{for } n = 1, 2, \dots,$$

we have

$$0 \leq \frac{1}{n} \leq \frac{1}{\sqrt{n}}, \quad \text{for } n = 1, 2, \dots$$

Since $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent (by Worked Exercise D32), we deduce from

part (b) of the Comparison Test that $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ is divergent.

Next we consider the series

$$\sum_{n=1}^{\infty} \frac{1}{2\sqrt{n} + 1}. \tag{4}$$

This seems likely to be divergent, because its terms are somewhat similar to those of $\sum_{n=1}^{\infty} (1/\sqrt{n})$, which we showed to be divergent in Worked Exercise D35. But how can we prove this?

We cannot use the Comparison Test for these two series because the inequality in step 2 of Strategy D10(b) does not hold. We *could* use the Comparison Test to compare series (4) with $\sum_{n=1}^{\infty} (1/n)$, in which case we would find that the inequality in step 2 holds eventually. However, the following useful result enables us to deduce the divergence of series (4) directly from the divergence of $\sum_{n=1}^{\infty} (1/\sqrt{n})$.

Theorem D31 Limit Comparison Test

Suppose that $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ have positive terms and that

$$\frac{a_n}{b_n} \rightarrow L \quad \text{as } n \rightarrow \infty,$$

where $L \neq 0$.

- (a) If $\sum_{n=1}^{\infty} b_n$ is convergent, then $\sum_{n=1}^{\infty} a_n$ is convergent.

- (b) If $\sum_{n=1}^{\infty} b_n$ is divergent, then $\sum_{n=1}^{\infty} a_n$ is divergent.

Proof (a) We know that the sequence (a_n/b_n) is convergent, so it must be bounded by Theorem D14 of Unit D2. Thus there is a positive constant K such that

$$\frac{a_n}{b_n} \leq K, \quad \text{for } n = 1, 2, \dots,$$

so

$$a_n \leq Kb_n, \quad \text{for } n = 1, 2, \dots.$$

We also know that $\sum_{n=1}^{\infty} b_n$ is convergent, so $\sum_{n=1}^{\infty} Kb_n$ is convergent, by the Multiple Rule. Hence, by part (a) of the Comparison Test, $\sum_{n=1}^{\infty} a_n$ is convergent.

- (b) As we saw in the proof of part (b) of the Comparison Test, the statement

$$\text{if } \sum_{n=1}^{\infty} b_n \text{ is divergent, then } \sum_{n=1}^{\infty} a_n \text{ is divergent}$$

is the contrapositive of the statement

$$\text{if } \sum_{n=1}^{\infty} a_n \text{ is convergent, then } \sum_{n=1}^{\infty} b_n \text{ is convergent,}$$

so these two statements are equivalent. Now since $L \neq 0$, we have

$$\frac{b_n}{a_n} \rightarrow \frac{1}{L} \text{ as } n \rightarrow \infty,$$

by the Quotient Rule for sequences from Unit D2. Thus the second statement follows from part (a) with the roles of a_n and b_n interchanged, which proves part (b). ■

Remarks

- The hypothesis $a_n/b_n \rightarrow L$ can be interpreted as saying that ' a_n behaves rather like b_n for large n '.
- The assumption that $L \neq 0$ is not needed in the proof of part (a), but it is essential in the proof of part (b).

As we did with the Comparison Test, we can formulate a convenient way to use the Limit Comparison Test via a 'guess then check' approach.

Strategy D11

- (a) To show that a series $\sum_{n=1}^{\infty} a_n$ of positive terms is *convergent* using the Limit Comparison Test, do the following.
1. Guess that $\sum_{n=1}^{\infty} a_n$ behaves like a comparable convergent series $\sum_{n=1}^{\infty} b_n$ of positive terms.
 2. Check that $\frac{a_n}{b_n} \rightarrow L \neq 0$ as $n \rightarrow \infty$ and deduce that $\sum_{n=1}^{\infty} a_n$ converges.
- (b) To show that a series $\sum_{n=1}^{\infty} a_n$ of positive terms is *divergent* using the Limit Comparison Test, do the following.
1. Guess that $\sum_{n=1}^{\infty} a_n$ behaves like a comparable divergent series $\sum_{n=1}^{\infty} b_n$ of positive terms.
 2. Check that $\frac{a_n}{b_n} \rightarrow L \neq 0$ as $n \rightarrow \infty$ and deduce that $\sum_{n=1}^{\infty} a_n$ diverges.

In either case, the first step is to find a suitable series $\sum_{n=1}^{\infty} b_n$ of positive terms to compare with the series $\sum_{n=1}^{\infty} a_n$ that we are investigating. To do this, we choose a series that we know to be convergent or divergent, and whose n th term b_n seems likely to behave in a similar way to a_n for large values of n . Carrying out the check in step 2 of Strategy D11 will then show whether or not our guess of $\sum_{n=1}^{\infty} b_n$ was suitable. This is illustrated in the next worked exercise.

Worked Exercise D36

Determine whether each of the following series converges or diverges.

$$(a) \sum_{n=1}^{\infty} \frac{1}{2\sqrt{n} + 1} \quad (b) \sum_{n=1}^{\infty} \frac{n+5}{3n^4 - n}$$

Solution

- (a) Following the discussion after Worked Exercise D35, we guess that terms of this series behave like $1/\sqrt{n}$ for large n . We know that $\sum_{n=1}^{\infty} (1/\sqrt{n})$ is divergent.

We use the Limit Comparison Test with

$$a_n = \frac{1}{2\sqrt{n} + 1}, \quad \text{for } n = 1, 2, \dots$$

and

$$b_n = \frac{1}{\sqrt{n}}, \quad \text{for } n = 1, 2, \dots$$

Both a_n and b_n are positive, and

$$\begin{aligned}\frac{a_n}{b_n} &= \frac{1}{2\sqrt{n}+1} \Big/ \frac{1}{\sqrt{n}} \\ &= \frac{\sqrt{n}}{2\sqrt{n}+1} \\ &= \frac{1}{2+1/\sqrt{n}} \rightarrow \frac{1}{2} \neq 0.\end{aligned}$$

💡 This follows from the Combination Rules for sequences since $(1/\sqrt{n})$ is a basic null sequence. 💡

Since $\sum_{n=1}^{\infty} (1/\sqrt{n})$ is divergent, it follows from part (b) of the

Limit Comparison Test that $\sum_{n=1}^{\infty} \frac{1}{2\sqrt{n}+1}$ is divergent.

- (b) 💡 For large n , the expression $\frac{n+5}{3n^4-n}$ is approximately $\frac{n}{3n^4} = \frac{1}{3n^3}$, so we guess that the terms of this series behave rather like $1/n^3$ for large n . We know that the series $\sum_{n=1}^{\infty} (1/n^3)$ is convergent. 💡

We use the Limit Comparison Test with

$$a_n = \frac{n+5}{3n^4-n}, \quad \text{for } n = 1, 2, \dots$$

and

$$b_n = \frac{1}{n^3}, \quad \text{for } n = 1, 2, \dots$$

Both a_n and b_n are positive, and

$$\begin{aligned}\frac{a_n}{b_n} &= \frac{n+5}{3n^4-n} \times \frac{n^3}{1} \\ &= \frac{n^4 + 5n^3}{3n^4 - n} \\ &= \frac{1 + 5/n}{3 - 1/n^3} \rightarrow \frac{1}{3} \neq 0.\end{aligned}$$

💡 This follows from the Combination Rules for sequences since $(1/n)$ and $(1/n^3)$ are basic null sequences. 💡

Since $\sum_{n=1}^{\infty} (1/n^3)$ is convergent, it follows from part (a) of the

Limit Comparison Test that $\sum_{n=1}^{\infty} \frac{n+5}{3n^4-n}$ is convergent.

In the next exercise you can practise using the Comparison Test and the Limit Comparison Test. If you can see a direct comparison with a series which you know to be convergent or divergent, then you can use the Comparison Test. If the terms of the series just behave like the terms of a known series for large values of n , then you should use the Limit Comparison Test.

Exercise D47

Use the Comparison Test or the Limit Comparison Test to determine whether each of the following series converges or diverges.

- (a) $\sum_{n=1}^{\infty} \frac{1}{n^3 + n}$
- (b) $\sum_{n=1}^{\infty} \frac{1}{n + \sqrt{n}}$
- (c) $\sum_{n=1}^{\infty} \frac{n+4}{2n^3 - n + 1}$
- (d) $\sum_{n=1}^{\infty} \frac{\cos^2(2n)}{n^3}$

Our next test for convergence is motivated in part by the properties of geometric series. Recall that the geometric series

$$a + ar + ar^2 + \cdots = \sum_{n=1}^{\infty} ar^n, \quad \text{where } a \neq 0,$$

is convergent if $|r| < 1$ but divergent if $|r| \geq 1$.

Theorem D32 Ratio Test

Suppose that $\sum_{n=1}^{\infty} a_n$ has positive terms and that $\frac{a_{n+1}}{a_n} \rightarrow l$ as $n \rightarrow \infty$.

- (a) If $0 \leq l < 1$, then $\sum_{n=1}^{\infty} a_n$ is convergent.
- (b) If $l > 1$ or $l = \infty$, then $\sum_{n=1}^{\infty} a_n$ is divergent.

Proof (a) We know that $0 \leq l < 1$, so we can choose $\varepsilon > 0$ such that $l + \varepsilon < 1$.

(For example, we can take $\varepsilon = \frac{1}{2}(1 - l)$.) Let $r = l + \varepsilon$. Since $r > l$, there is a positive integer N such that

$$\frac{a_{n+1}}{a_n} \leq r, \quad \text{for all } n \geq N.$$

This is illustrated in Figure 6. Thus, for $n \geq N$, we have

$$\frac{a_n}{a_N} = \left(\frac{a_n}{a_{n-1}} \right) \left(\frac{a_{n-1}}{a_{n-2}} \right) \cdots \left(\frac{a_{N+1}}{a_N} \right) \leq r^{n-N},$$

since each of the expressions in brackets is at most r . Hence

$$a_n \leq a_N r^{n-N}, \quad \text{for } n \geq N. \quad (5)$$

Now

$$\sum_{n=N}^{\infty} a_N r^{n-N} = a_N + a_N r + a_N r^2 + \cdots$$

is a geometric series with first term a_N and common ratio r . Since $0 < r < 1$, this series is convergent. Thus, by inequality (5) and the

Comparison Test, $\sum_{n=1}^{\infty} a_n$ is also convergent, as required.

(b) Since

$$\frac{a_{n+1}}{a_n} \rightarrow \infty \quad \text{or} \quad \frac{a_{n+1}}{a_n} \rightarrow l,$$

where $l > 1$, there is a positive integer N such that

$$\frac{a_{n+1}}{a_n} \geq 1, \quad \text{for all } n \geq N,$$

as shown in Figure 7. Thus, for $n \geq N$, we have

$$\frac{a_n}{a_N} = \left(\frac{a_n}{a_{n-1}} \right) \left(\frac{a_{n-1}}{a_{n-2}} \right) \cdots \left(\frac{a_{N+1}}{a_N} \right) \geq 1,$$

since each of the expressions in brackets is at least 1. Hence

$$a_n \geq a_N > 0, \quad \text{for } n \geq N,$$

so (a_n) cannot be a null sequence. It follows, by the Non-null Test,

that $\sum_{n=1}^{\infty} a_n$ is divergent. ■

Remarks

- With the Ratio Test, we concentrate on the series $\sum_{n=1}^{\infty} a_n$ itself and do not need to compare it with some other series $\sum_{n=1}^{\infty} b_n$.
- If $l = 1$, the Ratio Test gives us no information on whether the series converges. (That is, the Ratio Test is *inconclusive* if $l = 1$.)

For example, if $a_n = \frac{1}{n}$ then

$$\frac{a_{n+1}}{a_n} = \frac{n}{n+1} = \frac{1}{1+1/n} \rightarrow 1 \text{ as } n \rightarrow \infty$$

and we have seen (in Worked Exercise D32) that the series $\sum_{n=1}^{\infty} (1/n)$ diverges.

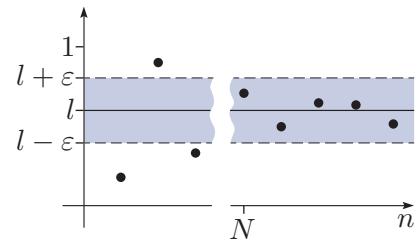


Figure 6 The sequence diagram for (a_{n+1}/a_n) if $0 \leq l < 1$

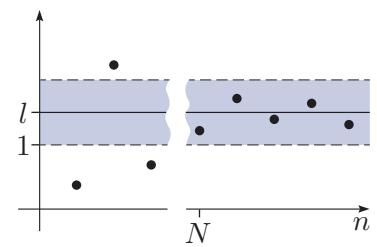


Figure 7 The sequence diagram for (a_{n+1}/a_n) if $l > 1$

On the other hand, if $a_n = \frac{1}{n^2}$ then

$$\frac{a_{n+1}}{a_n} = \frac{n^2}{(n+1)^2} = \frac{1}{1 + 2/n + 1/n^2} \rightarrow 1 \text{ as } n \rightarrow \infty,$$

but as we have seen (in Worked Exercise D33), the series $\sum_{n=1}^{\infty} (1/n^2)$ converges.

3. When using the Ratio Test, we obtain a_{n+1} by replacing each instance of n by $n + 1$ in the formula for a_n .



Jean le Rond d'Alembert

The Ratio Test was first published by the French mathematician Jean le Rond d'Alembert (1717–1783) in 1768. D'Alembert's interest in mechanics led him to take a Newtonian view of the foundations of the calculus, and he was among the first to regard the method of limits as fundamental. He was the editor of the mathematical and scientific articles in the *Encyclopédie* – a remarkable series of volumes which constitutes one of the major documents of the Enlightenment – and wrote many of the articles himself, including the ones on *Differential* and *Limite*.

Worked Exercise D37

Use the Ratio Test to determine whether the following series are convergent.

$$(a) \sum_{n=1}^{\infty} \frac{n}{2^n} \quad (b) \sum_{n=1}^{\infty} \frac{10^n}{n!}$$

Solution

(a) Let

$$a_n = \frac{n}{2^n}, \quad \text{for } n = 1, 2, \dots$$

Then a_n is positive, and

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{n+1}{2^{n+1}} \times \frac{2^n}{n} \\ &= \frac{n+1}{2n} \\ &= \frac{1 + 1/n}{2} \rightarrow \frac{1}{2} \text{ as } n \rightarrow \infty. \end{aligned}$$

Since $0 < \frac{1}{2} < 1$, it follows from the Ratio Test that $\sum_{n=1}^{\infty} \frac{n}{2^n}$ is convergent.

(b) Let

$$a_n = \frac{10^n}{n!}, \quad \text{for } n = 1, 2, \dots$$

Then a_n is positive, and

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \frac{10^{n+1}}{(n+1)!} \times \frac{n!}{10^n} \\ &= \frac{10}{n+1} \rightarrow 0 \text{ as } n \rightarrow \infty.\end{aligned}$$

It follows from the Ratio Test that $\sum_{n=1}^{\infty} \frac{10^n}{n!}$ is convergent.

Exercise D48

Use the Ratio Test to determine whether the following series are convergent.

$$(a) \sum_{n=1}^{\infty} \frac{n^3}{n!} \quad (b) \sum_{n=1}^{\infty} \frac{n^2 2^n}{n!} \quad (c) \sum_{n=1}^{\infty} \frac{(2n)!}{n^n}$$

Hint: In part (c) you need to use the fact (from Unit D2) that $(1 + 1/n)^n \rightarrow e$ as $n \rightarrow \infty$.

2.2 Basic series

When studying sequences in Unit D2, we made great use of a library of basic sequences. You will now see that there is also a library of basic series whose convergence or divergence is known. We can determine the convergence or divergence of a large number of other series from these basic series by using our tests.

Theorem D33 Basic series

The following series are convergent:

- (a) $\sum_{n=1}^{\infty} \frac{1}{n^p}$, for $p \geq 2$
- (b) $\sum_{n=1}^{\infty} c^n$, for $0 \leq c < 1$
- (c) $\sum_{n=1}^{\infty} n^p c^n$, for $p > 0$, $0 \leq c < 1$
- (d) $\sum_{n=1}^{\infty} \frac{c^n}{n!}$, for $c \geq 0$.

The following series is divergent:

- (e) $\sum_{n=1}^{\infty} \frac{1}{n^p}$, for $0 < p \leq 1$.

Proof (a) This series is convergent, by the Comparison Test, since if $p \geq 2$, then

$$\frac{1}{n^p} \leq \frac{1}{n^2}, \quad \text{for } n = 1, 2, \dots,$$

and the series $\sum_{n=1}^{\infty} (1/n^2)$ is convergent.

(In Unit F3 *Integration* we will prove that $\sum_{n=1}^{\infty} (1/n^p)$ is convergent, for all $p > 1$.)

(b) The series $\sum_{n=1}^{\infty} c^n$ is a geometric series with common ratio c , so it converges if $0 \leq c < 1$.

(c) Let

$$a_n = n^p c^n, \quad n = 1, 2, \dots$$

Then put $b = \sqrt{c}$ and express a_n as

$$a_n = n^p (b \times b)^n = (n^p b^n) b^n, \quad \text{for } n = 1, 2, \dots \quad (6)$$

Now $0 \leq b < 1$, so $(n^p b^n)$ is a basic null sequence. Hence, for some positive integer N , we have

$$n^p b^n < 1, \quad \text{for } n > N,$$

and thus, by equation (6),

$$a_n < b^n, \quad \text{for } n > N.$$

But $\sum_{n=1}^{\infty} b^n$ is a convergent geometric series, so $\sum_{n=1}^{\infty} a_n$ is convergent by the Comparison Test.

(d) Let

$$a_n = \frac{c^n}{n!}, \quad n = 1, 2, \dots$$

If $c = 0$, then the series is clearly convergent. For $c \neq 0$,

$$\frac{a_{n+1}}{a_n} = \frac{c^{n+1}}{(n+1)!} \Big/ \frac{c^n}{n!} = \frac{c^{n+1}}{(n+1)!} \times \frac{n!}{c^n} = \frac{c}{n+1}.$$

Thus

$$\frac{a_{n+1}}{a_n} \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

and we deduce from the Ratio Test that $\sum_{n=1}^{\infty} \frac{c^n}{n!}$ is convergent.

(e) This series is divergent by the Comparison Test, since if $p \leq 1$, then

$$\frac{1}{n^p} \geq \frac{1}{n}, \quad \text{for } n = 1, 2, \dots$$

and $\sum_{n=1}^{\infty} (1/n)$ is divergent. ■

Exercise D49

Verify that the following are all basic series, stating their type according to the list in Theorem D33 and giving the values of any parameters p and c . Hence determine which of the series are convergent.

- (a) $\sum_{n=1}^{\infty} \frac{2^n}{n!}$
- (b) $\sum_{n=1}^{\infty} \frac{1}{n^{5/2}}$
- (c) $\sum_{n=1}^{\infty} \frac{1}{n^{2/3}}$
- (d) $\sum_{n=1}^{\infty} \frac{n}{2^n}$
- (e) $\sum_{n=1}^{\infty} \frac{1}{4^n}$

3 Series with positive and negative terms

The study of series $\sum a_n$ with $a_n \geq 0$ for all values of n is relatively straightforward because the sequence (s_n) of partial sums is increasing. Similarly, if $a_n \leq 0$ for all values of n , then (s_n) is decreasing.

It is harder to determine the behaviour of a series with both positive and negative terms because (s_n) is neither increasing nor decreasing. However, if the sequence (a_n) contains only finitely many negative terms, then the sequence (s_n) is *eventually* increasing, and we can apply the results of Section 2. Similarly, if (a_n) contains only finitely many positive terms, then the sequence (s_n) is *eventually* decreasing, and we can again apply the results of Section 2 to the series $\sum_{n=1}^{\infty} (-a_n)$. For example, the convergence of the series

$$1 + 2 + 3 - \frac{1}{4^2} - \frac{1}{5^2} - \frac{1}{6^2} - \dots$$

follows from that of $\sum_{n=1}^{\infty} (1/n^2)$, by the Multiple Rule with $\lambda = -1$.

In this section we look at series such as

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots$$

and

$$1 + \frac{1}{2^2} - \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} - \frac{1}{6^2} + \dots,$$

which contain infinitely many terms of either sign. For such series the partial sums alternately increase and decrease infinitely often (so the snake discussed in Subsection 1.1 grows and shrinks infinitely often!). We give two methods which can often be used to prove that such series are convergent.

3.1 Absolute convergence

Suppose that we want to determine the behaviour of the infinite series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \frac{1}{5^2} - \frac{1}{6^2} + \dots \quad (7)$$

We know that the series

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \dots \quad (8)$$

is convergent. Does this imply that series (7) is also convergent? In fact it does, as we now prove.

Consider the two related series

$$1 + 0 + \frac{1}{3^2} + 0 + \frac{1}{5^2} + 0 + \dots$$

and

$$0 + \frac{1}{2^2} + 0 + \frac{1}{4^2} + 0 + \frac{1}{6^2} + \dots$$

Each of these series contains only non-negative terms and is dominated by series (8), so they are both convergent, by the Comparison Test. Applying the Sum Rule for series, and the Multiple Rule for series with $\lambda = -1$, we deduce that the series

$$1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \frac{1}{5^2} - \frac{1}{6^2} + \dots$$

is convergent.

The type of argument just given is the basis for a concept called *absolute convergence*, which we now define.

Definition

The series $\sum_{n=1}^{\infty} a_n$ is **absolutely convergent** if $\sum_{n=1}^{\infty} |a_n|$ is convergent.

If the terms a_n are all non-negative, then *absolute convergence* and *convergence* have the same meaning.

The series (7) is absolutely convergent because the series $\sum_{n=1}^{\infty} (1/n^2)$ is convergent. However, the series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots \quad (9)$$

is not absolutely convergent because the series $\sum_{n=1}^{\infty} (1/n)$ is divergent.

As the name suggests, every absolutely convergent series is convergent.

Theorem D34 Absolute Convergence Test

If $\sum_{n=1}^{\infty} a_n$ is absolutely convergent, then $\sum_{n=1}^{\infty} |a_n|$ is convergent.

Proof We know that $\sum_{n=1}^{\infty} |a_n|$ is convergent, and we want to prove that $\sum_{n=1}^{\infty} a_n$ is convergent.

To do this, we define two new series $\sum_{n=1}^{\infty} b_n$ and $\sum_{n=1}^{\infty} c_n$, where

$$b_n = \begin{cases} a_n, & \text{if } a_n \geq 0, \\ 0, & \text{if } a_n < 0, \end{cases} \quad c_n = \begin{cases} 0, & \text{if } a_n \geq 0, \\ -a_n, & \text{if } a_n < 0. \end{cases}$$

Both the series $\sum_{n=1}^{\infty} b_n$ and $\sum_{n=1}^{\infty} c_n$ have non-negative terms, and

$$b_n \leq |a_n|, \quad \text{for } n = 1, 2, \dots, \tag{10}$$

and

$$c_n \leq |a_n|, \quad \text{for } n = 1, 2, \dots. \tag{11}$$

Since $\sum_{n=1}^{\infty} |a_n|$ is convergent, we deduce that $\sum_{n=1}^{\infty} b_n$ and $\sum_{n=1}^{\infty} c_n$ are convergent, by the Comparison Test. Thus

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} (b_n - c_n) = \sum_{n=1}^{\infty} b_n - \sum_{n=1}^{\infty} c_n \tag{12}$$

is convergent, by the Combination Rules for series. ■

Because $\sum_{n=1}^{\infty} (1/n^2)$ is convergent, it follows from the Absolute Convergence Test that series (7) is convergent, as we have already seen. Indeed, however we distribute the plus and minus signs amongst the terms of the sequence $(1/n^2)$, the resulting series is convergent.

However, the Absolute Convergence Test tells us nothing about the behaviour of series (9), nor about similar series such as

$$1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} + \frac{1}{8} - \frac{1}{9} + \dots. \tag{13}$$

The series $\sum_{n=1}^{\infty} (1/n)$ is divergent, so these two series are not absolutely convergent. You will see later how to use other methods to determine whether series (9) and (13) are convergent.

Worked Exercise D38

Prove that the following series are convergent.

$$(a) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^3} \quad (b) \sum_{n=1}^{\infty} \frac{\cos n}{2^n}$$

Solution

(a) Let

$$a_n = \frac{(-1)^{n+1}}{n^3}, \quad \text{for } n = 1, 2, \dots$$

Then

$$|a_n| = \frac{1}{n^3}, \quad \text{for } n = 1, 2, \dots$$

We know that $\sum_{n=1}^{\infty} (1/n^3)$ is convergent, so by the Absolute Convergence Test,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^3} \text{ is convergent.}$$

(b) Let

$$a_n = \frac{\cos n}{2^n}, \quad \text{for } n = 1, 2, \dots$$

 In this case we use the fact that $|\cos n| \leq 1$ to get an upper bound for the size of $|a_n|$. We can then use the Comparison Test to show that $\sum |a_n|$ is convergent. 

Then

$$|a_n| \leq \frac{1}{2^n}, \quad \text{for } n = 1, 2, \dots,$$

because $|\cos n| \leq 1$, for $n = 1, 2, \dots$

Since $\sum_{n=1}^{\infty} (1/2^n)$ is a basic convergent series, it follows from the Comparison Test that $\sum_{n=1}^{\infty} |a_n|$ is convergent. Hence, by the Absolute Convergence Test,

$$\sum_{n=1}^{\infty} \frac{\cos n}{2^n} \text{ is convergent.}$$

Exercise D50

Prove that the following series are convergent.

$$(a) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}n}{n^3 + 1} \quad (b) 1 + \frac{1}{2} - \frac{1}{4} + \frac{1}{8} + \frac{1}{16} - \frac{1}{32} + \dots$$

The Absolute Convergence Test states that if the series $\sum |a_n|$ is convergent, then $\sum a_n$ is also convergent. The following result relates the sums of these two convergent series. This result generalises the Triangle Inequality for n real numbers,

$$|a_1 + a_2 + \cdots + a_n| \leq |a_1| + |a_2| + \cdots + |a_n|,$$

that is,

$$\left| \sum_{k=1}^n a_k \right| \leq \sum_{k=1}^n |a_k|,$$

which you met in Subsection 3.1 of Unit D1 *Numbers*.

Theorem D35 Triangle Inequality (infinite form)

If $\sum_{n=1}^{\infty} a_n$ is absolutely convergent, then

$$\left| \sum_{n=1}^{\infty} a_n \right| \leq \sum_{n=1}^{\infty} |a_n|.$$

Proof We use the series $\sum_{n=1}^{\infty} b_n$ and $\sum_{n=1}^{\infty} c_n$, introduced in the proof of the Absolute Convergence Test. Since the numbers b_n and c_n are all non-negative, we obtain the following inequalities from equation (12):

$$-\sum_{n=1}^{\infty} c_n \leq \sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} b_n.$$

Thus, by inequalities (10) and (11), we deduce that

$$-\sum_{n=1}^{\infty} |a_n| \leq \sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} |a_n|,$$

which gives the required inequality

$$\left| \sum_{n=1}^{\infty} a_n \right| \leq \sum_{n=1}^{\infty} |a_n|. \quad \blacksquare$$

Exercise D51

Show that the series

$$\frac{1}{2} - \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \frac{1}{32} - \frac{1}{64} + \cdots$$

is convergent, and that its sum lies in the interval $[-1, 1]$.

(In this series the signs of the terms are $+, -, -,$ repeated infinitely often. You do not need to find the sum of the series.)

3.2 Alternating Test

Suppose that we want to determine the behaviour of the following infinite series, in which the terms have alternating signs:

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots \quad (14)$$

(You met this series earlier as series (9).) The Absolute Convergence Test does not help us with this series because $\sum_{n=1}^{\infty} (1/n)$ is divergent. In fact, series (14) is convergent. To see why, we first calculate some of the partial sums and plot them on a sequence diagram:

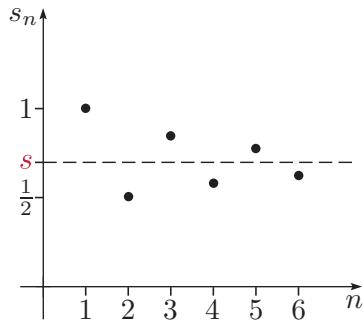


Figure 8 The sequence diagram for (s_n)

$$\begin{aligned}s_1 &= 1, \\s_2 &= 1 - \frac{1}{2} = 0.5, \\s_3 &= 1 - \frac{1}{2} + \frac{1}{3} = 0.8\bar{3}, \\s_4 &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} = 0.58\bar{3}, \\s_5 &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} = 0.78\bar{3}, \\s_6 &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} = 0.61\bar{6}.\end{aligned}$$

The sequence diagram is shown in Figure 8. This suggests that the odd subsequence (s_{2k-1}) is decreasing:

$$s_1 \geq s_3 \geq s_5 \geq \dots \geq s_{2k-1} \geq \dots,$$

and that the even subsequence (s_{2k}) is increasing:

$$s_2 \leq s_4 \leq s_6 \leq \dots \leq s_{2k} \leq \dots.$$

Also, the terms of (s_{2k-1}) all exceed the terms of (s_{2k}) , and both subsequences appear to converge to a common limit s , which lies between the odd and even partial sums.

To prove this, we write the even partial sum s_{2k} as

$$s_{2k} = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \dots + \left(\frac{1}{2k-1} - \frac{1}{2k}\right).$$

All the expressions in brackets are positive, so the subsequence (s_{2k}) is increasing.

We can also write

$$s_{2k} = 1 - \left(\frac{1}{2} - \frac{1}{3}\right) - \left(\frac{1}{4} - \frac{1}{5}\right) - \dots - \left(\frac{1}{2k-2} - \frac{1}{2k-1}\right) - \frac{1}{2k}.$$

Again, all the expressions in brackets are positive, so (s_{2k}) is bounded above by 1.

Hence, by the Monotone Convergence Theorem,

$$\lim_{k \rightarrow \infty} s_{2k} = s,$$

for some s . Since

$$s_{2k} = s_{2k-1} - \frac{1}{2k}, \quad \text{for } k = 1, 2, \dots,$$

and the sequence $(1/2k)$ is null, we deduce that

$$\lim_{k \rightarrow \infty} s_{2k-1} = \lim_{k \rightarrow \infty} \left(s_{2k} + \frac{1}{2k} \right) = s,$$

by the Sum Rule for sequences. Thus the odd and even subsequences of (s_n) both tend to the same limit s , so $s_n \rightarrow s$ as $n \rightarrow \infty$, by Theorem D21 in Unit D2. Hence, the series $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ is convergent, with sum s .

(In fact, $s = \log_e 2 \approx 0.69$ but we do not show this here.)

The same method can be used to prove the following general result which is also called the Leibniz Test.

Theorem D36 Alternating Test

Let

$$a_n = (-1)^{n+1} b_n, \quad n = 1, 2, \dots,$$

where (b_n) is a decreasing null sequence with positive terms. Then

$$\sum_{n=1}^{\infty} a_n = b_1 - b_2 + b_3 - b_4 + \dots \text{ is convergent.}$$

Proof We can write any even partial sum s_{2k} as

$$s_{2k} = (b_1 - b_2) + (b_3 - b_4) + \dots + (b_{2k-1} - b_{2k}).$$

Since the sequence (b_n) is decreasing, all the expressions in brackets are non-negative, so the even subsequence (s_{2k}) is increasing.

We can also write

$$s_{2k} = b_1 - (b_2 - b_3) - (b_4 - b_5) - \dots - (b_{2k-2} - b_{2k-1}) - b_{2k}.$$

Again, all the expressions in brackets are non-negative, so the subsequence (s_{2k}) is bounded above by b_1 .

Hence, by the Monotone Convergence Theorem,

$$\lim_{k \rightarrow \infty} s_{2k} = s,$$

for some s . Since

$$s_{2k} = s_{2k-1} - b_{2k}, \quad \text{for } k = 1, 2, \dots,$$

and the sequence (b_n) is null, we deduce that

$$\lim_{k \rightarrow \infty} s_{2k-1} = \lim_{k \rightarrow \infty} (s_{2k} + b_{2k}) = s,$$

by the Sum Rule for sequences. Thus the odd and even subsequences of (s_n) both tend to the same limit s . Hence, by Theorem D21 in Unit D2, s_n

tends to s , so $\sum_{n=1}^{\infty} a_n$ is convergent, with sum s . ■

When you apply the Alternating Test there are a number of conditions to check. We now describe these in the form of a strategy.

Strategy D12

To prove that $\sum_{n=1}^{\infty} a_n$ is convergent using the Alternating Test, check that

$$a_n = (-1)^{n+1} b_n, \quad n = 1, 2, \dots,$$

where

1. $b_n \geq 0$, for $n = 1, 2, \dots$
2. (b_n) is a null sequence
3. (b_n) is decreasing.

When you use this strategy and are checking that the sequence (b_n) is null, you may find it helpful to use the techniques you met in Unit D2, including the list of basic null sequences.

Here are some examples of the use of Strategy D12.

Worked Exercise D39

Prove that the following series are convergent.

$$(a) \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\sqrt{n}} \quad (b) \sum_{n=1}^{\infty} \frac{(-1)^{n+1} n}{2n^2 - 1}$$

Solution

- (a) Let

$$a_n = \frac{(-1)^{n+1}}{\sqrt{n}}, \quad \text{for } n = 1, 2, \dots$$

Then $a_n = (-1)^{n+1}b_n$, where

$$b_n = 1/\sqrt{n}, \quad \text{for } n = 1, 2, \dots$$

Now

1. $b_n \geq 0$, for $n = 1, 2, \dots$
2. (b_n) is a basic null sequence
3. (b_n) is decreasing, because $(1/b_n) = (\sqrt{n})$ is increasing.

Hence, by the Alternating Test,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\sqrt{n}} \text{ is convergent.}$$

(b) Let

$$a_n = \frac{(-1)^{n+1}n}{2n^2 - 1}, \quad \text{for } n = 1, 2, \dots$$

Then $a_n = (-1)^{n+1}b_n$, where

$$b_n = \frac{n}{2n^2 - 1}, \quad \text{for } n = 1, 2, \dots$$

Now

1. $b_n \geq 0$, for $n = 1, 2, \dots$
2. since

$$b_n = \frac{1/n}{2 - 1/n^2} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

(b_n) is null

3. (b_n) is decreasing, because

$$\left(\frac{1}{b_n}\right) = \left(\frac{2n^2 - 1}{n}\right) = \left(2n - \frac{1}{n}\right)$$

is increasing.

Hence, by the Alternating Test,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}n}{2n^2 - 1} \text{ is convergent.}$$

Exercise D52

Determine which of the following series are convergent.

- (a) $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^{1/3}}$ (b) $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n + n^{1/2}}$ (c) $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{n}{n+2}$

3.3 General strategy

We now give a strategy for applying the tests for convergence (or divergence) of a series. For each test, the strategy briefly indicates the circumstances under which that test can be used.

Strategy D13

To determine whether a series $\sum a_n$ is convergent or divergent, do the following.

1. If you think that the sequence of terms (a_n) is non-null, then try the **Non-null Test**.
2. If $\sum a_n$ has non-negative terms, then try one of these tests.

Basic series Is $\sum a_n$ a basic series, or a combination of these?

Comparison Test Is $a_n \leq b_n$, where $\sum b_n$ is convergent, or $a_n \geq b_n \geq 0$, where $\sum b_n$ is divergent?

Limit Comparison Test Does a_n behave like b_n for large n (that is, does $a_n/b_n \rightarrow L \neq 0$), where $\sum b_n$ is a series that you know converges or diverges?

Ratio Test Does $a_{n+1}/a_n \rightarrow l \neq 1$?

3. If $\sum a_n$ has infinitely many positive and negative terms, then try one of these tests.

Absolute Convergence Test Is $\sum |a_n|$ convergent?
(Use step 2.)

Alternating Test Is $a_n = (-1)^{n+1}b_n$, where (b_n) is non-negative, null and decreasing?

Remarks

1. When applying these tests, you do not need to try to prove that the sequence (a_n) is null (though in the case of the Alternating Test, you do need to show that the sequence (b_n) is null).
2. If the terms a_n of the series $\sum a_n$ are non-positive, apply step 2 of the strategy to the series $\sum(-a_n)$ and then use the Multiple Rule with $\lambda = -1$.
3. If none of steps 1–3 of the strategy gives a result, then you could try using first principles by working directly with the sequence (s_n) of partial sums.

4. The following suggestions may also be helpful.

- If a_n is positive and includes $n!$ or c^n , then consider the Ratio Test.
- If a_n is positive and has dominant term n^p , then consider the Comparison Test or the Limit Comparison Test.
- If a_n includes a sine or cosine term, then use the fact that this term is bounded and consider the Comparison Test and the Absolute Convergence Test.

Exercise D53

Determine which of the following series are convergent.

$$\begin{array}{lll} \text{(a)} \sum_{n=1}^{\infty} \frac{5n+2^n}{3^n} & \text{(b)} \sum_{n=1}^{\infty} \frac{3}{2n^3-1} & \text{(c)} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n \log(n+1)} \\ \text{(d)} \sum_{n=1}^{\infty} \frac{(-1)^{n+1} n^2}{n^2+1} & \text{(e)} \sum_{n=1}^{\infty} \frac{(-1)^{n+1} n}{n^3+5} & \text{(f)} \sum_{n=1}^{\infty} \frac{2^n}{n^6} \end{array}$$

4 The exponential function

In this section you will see how e^x can be represented as an infinite series of powers of x . This representation is then used to prove that the number e is irrational, and also that $e^x e^y = e^{x+y}$ for any real numbers x and y . This section is not assessed and is for reading only.

4.1 Definition of e^x

In Subsection 5.3 of Unit D2 we defined $e = 2.71828\dots$ to be the limit

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

We also stated that if $x > 0$, then

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n,$$

We can now use infinite series to give an alternative definition of e^x .

Figure 9 shows graphs of the first four partial sum functions

$$s_0(x) = 1, \quad s_1(x) = 1 + x, \quad s_2(x) = 1 + x + \frac{x^2}{2!}, \quad \dots, \quad$$

of the following infinite series of powers of x :

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots. \quad (15)$$

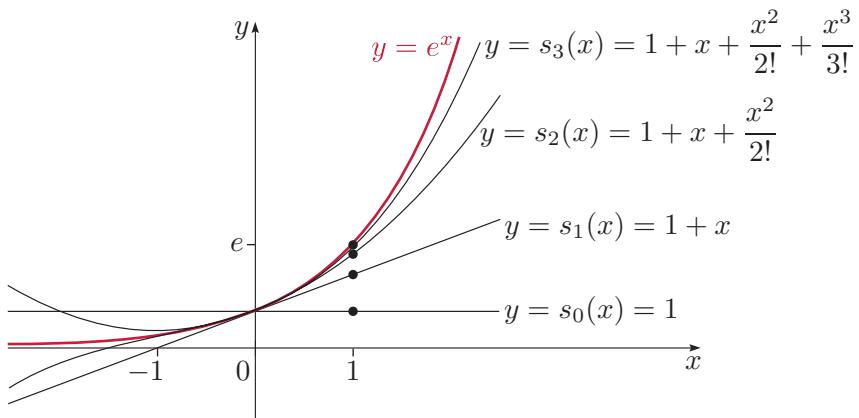


Figure 9 The partial sum functions of $\sum_{n=0}^{\infty} (x^n/n!)$

We know that series (15) is convergent for all real numbers x since it is a basic convergent series of type (a). As the sum of the series depends on x , it defines a function of x . If you test different values, then you will find that this sum function appears to be e^x . In particular, when $x = 1$, the sum of the series

$$\sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots$$

is approximately 2.718...

It can be proved that series (15) does converge to e^x for all $x \in \mathbb{R}$ and we give this result, for $x \geq 0$, as our next theorem. If you are short of time, then skim read the proof, noting the main points.

Theorem D37

If $x \geq 0$, then

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

Proof We give the proof only for the case $x = 1$; the general case is similar. We have to show that

$$\sum_{n=0}^{\infty} \frac{1}{n!} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e.$$

The n th partial sum of this convergent series is

$$s_n = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!}, \quad \text{so} \quad \lim_{n \rightarrow \infty} s_n = \sum_{n=0}^{\infty} \frac{1}{n!}.$$

Now, by the Binomial Theorem, we have

$$\left(1 + \frac{1}{n}\right)^n = 1 + n\left(\frac{1}{n}\right) + \frac{n(n-1)}{2!}\left(\frac{1}{n}\right)^2 + \cdots + \left(\frac{1}{n}\right)^n. \quad (16)$$

The general term in this expansion is of the form

$$\begin{aligned} & \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} \left(\frac{1}{n}\right)^k \\ &= \frac{1}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right), \end{aligned} \quad (17)$$

where $0 \leq k \leq n$. This last product is at most $1/k!$, since each expression in brackets is at most 1, so

$$\left(1 + \frac{1}{n}\right)^n \leq 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!} = s_n,$$

by equation (16). Thus, by the Limit Inequality Rule for sequences,

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \leq \lim_{n \rightarrow \infty} s_n = \sum_{n=0}^{\infty} \frac{1}{n!}. \quad (18)$$

On the other hand, if $0 \leq m \leq n$, then (by equations (16) and (17))

$$\begin{aligned} \left(1 + \frac{1}{n}\right)^n &\geq 1 + 1 + \frac{1}{2!} \left(1 - \frac{1}{n}\right) + \cdots \\ &\quad \cdots + \frac{1}{m!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{m-1}{n}\right). \end{aligned}$$

Keeping m fixed and taking limits of the above inequality as $n \rightarrow \infty$, we obtain

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \geq 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{m!} = s_m,$$

by the Limit Inequality Rule. Applying this rule once more to $e \geq s_m$ and using the fact that e is a constant, we obtain

$$e \geq \lim_{m \rightarrow \infty} s_m = \sum_{m=0}^{\infty} \frac{1}{m!} = \sum_{n=0}^{\infty} \frac{1}{n!}. \quad (19)$$

Combining inequalities (18) and (19) gives $\sum_{n=0}^{\infty} \frac{1}{n!} = e$. ■

By Theorem D37, we can use the series $\sum_{n=0}^{\infty} (x^n/n!)$ as an alternative definition of e^x for all $x \geq 0$. We then define e^x for $x < 0$ as the reciprocal of e^{-x} . For example, $e^{-\pi} = (e^{\pi})^{-1} = 1/e^{\pi}$.

Definition

For $x \geq 0$,

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

For $x < 0$,

$$e^x = (e^{-x})^{-1}.$$

Remarks

1. The equations

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

are also true if x is negative, but we shall not prove this here. The fact that

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad \text{for } x < 0,$$

is proved in Unit F4 *Power series*.

2. The exponential function $x \mapsto e^x$ is often called **exp** and we may write

$$\begin{aligned} \exp : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto e^x. \end{aligned}$$

4.2 Calculating e

The representation of e by the infinite series

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}$$

provides a more efficient method of calculating approximate values for e than the equation $e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$. This is illustrated by the following table of approximate values for $e = 2.718\,281\,828\,45\dots$

n	1	2	3	4	5
$(1 + 1/n)^n$	2	2.25	2.37	2.44	2.49
$\sum_{k=0}^n \frac{1}{k!}$	2	2.50	2.67	2.71	2.717

We now estimate how quickly the sequence of partial sums

$$s_n = 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!}, \quad n = 1, 2, \dots,$$

converges to e . The difference between e and s_n is given by

$$\begin{aligned} e - s_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \cdots \\ &= \frac{1}{(n+1)!} \left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \cdots \right) \\ &< \frac{1}{(n+1)!} \left(1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \cdots \right). \end{aligned}$$

The inequality above holds because each term inside the large brackets has been replaced by a larger term. The resulting expression is a geometric series with first term 1 and common ratio $1/(n+1)$, so its sum is

$$\frac{1}{1 - 1/(n+1)} = \frac{n+1}{n}.$$

Hence

$$0 < e - s_n < \frac{1}{(n+1)!} \times \frac{n+1}{n} = \frac{1}{n!} \times \frac{1}{n}, \quad \text{for } n = 1, 2, \dots \quad (20)$$

Thus the difference between e and s_n is extremely small when n is large.

Inequality (20) can also be used to show that e is irrational.

Theorem D38

The number e is irrational.

Proof We use proof by contradiction. Suppose that $e = m/n$, where m and n are positive integers. Then, by inequality (20), for this particular integer n we have

$$0 < e - s_n < \frac{1}{n!} \times \frac{1}{n},$$

so

$$0 < n!(e - s_n) < \frac{1}{n}.$$

Since we are assuming that $e = m/n$, we have

$$0 < n! \left(\frac{m}{n} - \left(1 + \frac{1}{2!} + \cdots + \frac{1}{n!} \right) \right) < \frac{1}{n}.$$

But the middle expression in this pair of inequalities is an integer, as you can check by multiplying it out, so we have found an integer which lies strictly between 0 and 1, since $1/n \leq 1$ for $n = 1, 2, \dots$. This is impossible, so e is not rational. ■



Joseph Fourier

Leonhard Euler in 1737 was the first to prove that e is irrational, although his proof, which used continued fractions, was not published until 1744. The first proof by contradiction is due to the French mathematician Joseph Fourier (1768–1830). It appeared in a collection of mathematical results published in 1815 by Fourier's compatriot, Janot de Stainville (1783–1828).

4.3 A fundamental property of e^x

We complete this section by showing that the function $f(x) = e^x$ satisfies one of the Index Laws which we stated in Unit D1. If you are short of time, then you may prefer to skim read the proof of this result.

Theorem D39

For any real numbers x and y , we have $e^{x+y} = e^x e^y$.

Proof First we prove the special case where x and y are both positive. We have

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

and

$$e^y = 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \dots$$

The following table contains some of the terms which occur when we multiply together partial sums of the power series for e^x and e^y .

	1	y	$\frac{y^2}{2!}$	$\frac{y^3}{3!}$...
1	1	y	$\frac{y^2}{2!}$	$\frac{y^3}{3!}$...
x	x	xy	$\frac{xy^2}{2!}$	$\frac{xy^3}{3!}$...
$\frac{x^2}{2!}$	$\frac{x^2}{2!}$	$\frac{x^2y}{2!}$	$\frac{x^2y^2}{2!2!}$	$\frac{x^2y^3}{2!3!}$...
$\frac{x^3}{3!}$	$\frac{x^3}{3!}$	$\frac{x^3y}{3!}$	$\frac{x^3y^2}{3!2!}$	$\frac{x^3y^3}{3!3!}$...
\vdots	\vdots	\vdots	\vdots	\vdots	

Adding the terms on the ‘lower left to upper right’ diagonals of the table gives:

1

$x + y$

$$\frac{x^2}{2!} + xy + \frac{y^2}{2!} = \frac{(x+y)^2}{2!}$$

$$\frac{x^3}{3!} + \frac{x^2y}{2!} + \frac{xy^2}{2!} + \frac{y^3}{3!} = \frac{(x+y)^3}{3!}$$

⋮

$$\frac{x^n}{n!} + \frac{x^{n-1}y}{(n-1)!} + \cdots + \frac{xy^{n-1}}{(n-1)!} + \frac{y^n}{n!} = \frac{(x+y)^n}{n!}$$

For any positive integer n , the product

$$\left(\sum_{k=0}^n \frac{x^k}{k!} \right) \left(\sum_{k=0}^n \frac{y^k}{k!} \right) = \left(1 + x + \cdots + \frac{x^n}{n!} \right) \left(1 + y + \cdots + \frac{y^n}{n!} \right)$$

includes all the terms in the first $n+1$ diagonals of the table up to the diagonal beginning $\frac{x^n}{n!}$; moreover, all the terms of the product are included

in the first $2n+1$ diagonals up to the diagonal containing $\frac{x^n y^n}{n! n!}$, which

begins $\frac{x^{2n}}{(2n)!}$.

Since x and y are non-negative, it follows that

$$\sum_{k=0}^n \frac{(x+y)^k}{k!} \leq \left(\sum_{k=0}^n \frac{x^k}{k!} \right) \left(\sum_{k=0}^n \frac{y^k}{k!} \right) \leq \sum_{k=0}^{2n} \frac{(x+y)^k}{k!}.$$

But

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{(x+y)^k}{k!} = e^{x+y} \quad \text{and} \quad \lim_{n \rightarrow \infty} \sum_{k=0}^{2n} \frac{(x+y)^k}{k!} = e^{x+y}.$$

Thus, by the Squeeze Rule and the Product Rule, we deduce that

$$e^x e^y = \lim_{n \rightarrow \infty} \left(\sum_{k=0}^n \frac{x^k}{k!} \right) \left(\sum_{k=0}^n \frac{y^k}{k!} \right) = e^{x+y},$$

as required.

If x and y are not both positive, then we can verify the equation

$$e^x e^y = e^{x+y}$$

by rearranging it so that all the powers are positive (using $e^x = (e^{-x})^{-1}$) and applying the special case just proved. For example, if $x > y > 0$, then

$$e^x e^{-y} = e^{x-y}$$

is equivalent to

$$e^x = e^{x-y} e^y,$$

which is true, since $x - y > 0$ and $y > 0$. ■

Summary

In this unit you have studied infinite series of the form $\sum_{n=1}^{\infty} a_n$. You have seen that such a series is said to be convergent if the sequence of partial sums (s_n) defined by

$$s_n = a_1 + a_2 + \cdots + a_n$$

is convergent, and is said to be divergent otherwise. You have also learnt that it is important to distinguish carefully between the sequence (s_n) and the sequence (a_n) . The Non-null Test tells you that, if a series is convergent, then the corresponding sequence (a_n) must be null; but if you know that the sequence (a_n) is null, then this gives you no information about whether or not the series is convergent.

You have met a number of tests that can be used to determine whether a series is convergent and a strategy to help you to work out which test is the most appropriate. If all the terms of a series are non-negative, then you may be able to use the Comparison Test or the Limit Comparison Test to compare with a basic series you know to be convergent or divergent; or you may be able to use the Ratio Test. For a series with infinitely many positive and negative terms, you may be able to use the Absolute Convergence Test or the Alternating Test.

Finally, you have seen how we can use a particular series to define the exponential function e^x , enabling us to prove that e is irrational and that $e^{x+y} = e^x e^y$. In Book F *Analysis 2* you will see how a wide range of functions can be defined using series when you study power series and their properties.

Learning outcomes

After working through this unit, you should be able to:

- explain what is meant by a *convergent* series $\sum_{n=1}^{\infty} a_n$
- write down the sum of a convergent *geometric* series
- find the sums of certain *telescoping* series
- use the Combination Rules for convergent series
- use the Non-null Test to recognise certain *divergent* series
- explain why $\sum_{n=1}^{\infty} (1/n)$ is divergent and $\sum_{n=1}^{\infty} (1/n^2)$ is convergent
- use the Comparison Test, the Limit Comparison Test and the Ratio Test
- recognise and use *basic* series
- explain the term *absolutely convergent* and use the Absolute Convergence Test
- use the Alternating Test
- use the given general strategy for determining whether a series is convergent or divergent
- appreciate that there are two equivalent definitions of e^x
- understand how the series definition of e^x enables us to prove that e is irrational, and that $e^{x+y} = e^x e^y$.

Solutions to exercises

Solution to Exercise D41

(a) Using the formula for summing a finite geometric series, with $a = r = -\frac{1}{3}$, we obtain

$$\begin{aligned}s_n &= \left(-\frac{1}{3}\right) + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^3 + \cdots + \left(-\frac{1}{3}\right)^n \\&= \frac{\left(-\frac{1}{3}\right)\left(1 - \left(-\frac{1}{3}\right)^n\right)}{1 - \left(-\frac{1}{3}\right)} \\&= -\frac{1}{4}\left(1 - \left(-\frac{1}{3}\right)^n\right).\end{aligned}$$

Since $\left(\left(-\frac{1}{3}\right)^n\right)$ is a basic null sequence,

$$\lim_{n \rightarrow \infty} s_n = -\frac{1}{4},$$

so

$$\sum_{n=1}^{\infty} \left(-\frac{1}{3}\right)^n \text{ is convergent, with sum } -\frac{1}{4}.$$

(b) In this case,

$$\begin{aligned}s_n &= (-1) + (-1)^2 + (-1)^3 + \cdots + (-1)^n \\&= -1 + 1 - 1 + \cdots + (-1)^n.\end{aligned}$$

Thus

$$s_n = \begin{cases} -1, & \text{if } n \text{ is odd,} \\ 0, & \text{if } n \text{ is even.} \end{cases}$$

Hence the odd subsequence (s_{2k-1}) converges to -1 and the even subsequence (s_{2k}) converges to 0 . Thus (s_n) has two convergent subsequences with different limits, so by the First Subsequence Rule (see Unit D2), (s_n) is divergent. It follows that the series $\sum_{n=1}^{\infty} (-1)^n$ is divergent.

(c) In this case, the first term in the series is a_0 , so the n th partial sum is the sum of $n+1$ terms. Using the formula for the sum of a finite geometric series with $a = 1$ and $r = \frac{1}{2}$, and summing $n+1$ terms, we obtain

$$\begin{aligned}s_n &= 1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \cdots + \left(\frac{1}{2}\right)^n \\&= \frac{1\left(1 - \left(\frac{1}{2}\right)^{n+1}\right)}{1 - \frac{1}{2}} \\&= 2\left(1 - \left(\frac{1}{2}\right)^{n+1}\right) = 2 - \left(\frac{1}{2}\right)^n.\end{aligned}$$

Since $\left(\left(\frac{1}{2}\right)^n\right)$ is a basic null sequence,

$$\lim_{n \rightarrow \infty} s_n = 2,$$

so

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n \text{ is convergent, with sum 2.}$$

You might have anticipated this result, since we proved at the beginning of Subsection 1.1 that

$$\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n = 1.$$

Thus, using the fact that $\left(\frac{1}{2}\right)^0 = 1$, we have

$$\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 1 + \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n = 2.$$

Solution to Exercise D42

We have

$$s_1 = \frac{1}{1 \times 2} = \frac{1}{2},$$

$$s_2 = \frac{1}{1 \times 2} + \frac{1}{2 \times 3} = \frac{1}{2} + \frac{1}{6} = \frac{2}{3},$$

$$s_3 = \frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \frac{1}{3 \times 4} = \frac{2}{3} + \frac{1}{12} = \frac{3}{4},$$

$$s_4 = \frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \frac{1}{3 \times 4} + \frac{1}{4 \times 5} = \frac{3}{4} + \frac{1}{20} = \frac{4}{5}.$$

Solution to Exercise D43

Using the given identity, we see that

$$s_n = \sum_{k=1}^n \frac{1}{k(k+2)} = \frac{1}{2} \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+2}\right).$$

Thus

$$\begin{aligned}s_n &= \frac{1}{1 \times 3} + \frac{1}{2 \times 4} + \frac{1}{3 \times 5} + \cdots + \frac{1}{n(n+2)} \\&= \frac{1}{2} \left(\left(1 - \frac{1}{3}\right) + \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{5}\right) \right. \\&\quad \left. + \left(\frac{1}{4} - \frac{1}{6}\right) + \left(\frac{1}{5} - \frac{1}{7}\right) + \cdots \right. \\&\quad \left. + \left(\frac{1}{n-1} - \frac{1}{n+1}\right) + \left(\frac{1}{n} - \frac{1}{n+2}\right) \right).\end{aligned}$$

All the terms in alternate brackets cancel, except for the first term in each of the first two brackets and the second term in each of the last two brackets. Thus

$$s_n = \frac{1}{2} \left(1 + \frac{1}{2} - \frac{1}{n+1} - \frac{1}{n+2} \right).$$

Since $\left(\frac{1}{n+1}\right)$ and $\left(\frac{1}{n+2}\right)$ are null sequences, it follows that

$$\lim_{n \rightarrow \infty} s_n = \frac{1}{2} \left(1 + \frac{1}{2} \right) = \frac{3}{4},$$

so

$$\sum_{n=1}^{\infty} \frac{1}{n(n+2)} \text{ is convergent, with sum } \frac{3}{4}.$$

Solution to Exercise D44

The series $\sum_{n=1}^{\infty} \left(-\frac{3}{4}\right)^n$ is a geometric series with $a = -\frac{3}{4}$ and $r = -\frac{3}{4}$. Since $-\frac{3}{4} < 1$, the series is convergent, with sum

$$\frac{-\frac{3}{4}}{1 - \left(-\frac{3}{4}\right)} = -\frac{3}{7}.$$

The series $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ is convergent, with sum 1;

see Subsection 1.2. Hence, by the Combination Rules,

$$\sum_{n=1}^{\infty} \left(\left(-\frac{3}{4}\right)^n - \frac{2}{n(n+1)} \right) \text{ is convergent,}$$

with sum $-\frac{3}{7} - (2 \times 1) = -\frac{17}{7}$.

Solution to Exercise D45

Let $a_n = \frac{(-1)^{n+1} n^2}{2n^2 + 1}$, so that $|a_n| = \frac{n^2}{2n^2 + 1}$.

By the Combination Rules for sequences,

$$\lim_{n \rightarrow \infty} \frac{n^2}{2n^2 + 1} = \lim_{n \rightarrow \infty} \frac{1}{2 + 1/n^2} = \frac{1}{2} \neq 0,$$

so the sequence (a_n) is not null.

Hence, by the Non-null Test,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1} n^2}{2n^2 + 1} \text{ is divergent.}$$

Solution to Exercise D46

Let

$$s_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}$$

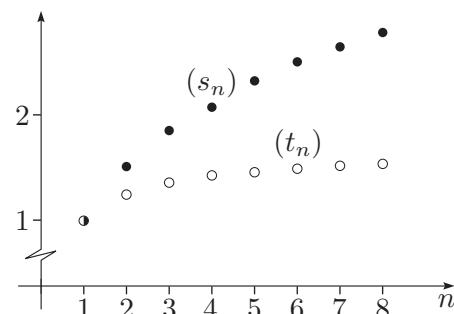
and

$$t_n = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots + \frac{1}{n^2}.$$

The values of the first eight partial sums are given below to two decimal places.

n	1	2	3	4	5	6	7	8
s_n	1	1.5	1.83	2.08	2.28	2.45	2.59	2.72
t_n	1	1.25	1.36	1.42	1.46	1.49	1.51	1.53

The sequences are plotted below.



Solution to Exercise D47

(a) We guess that this series is dominated by $\sum(1/n^3)$. We have

$$n^3 + n \geq n^3 \geq 0, \quad \text{for } n = 1, 2, \dots,$$

so

$$0 \leq \frac{1}{n^3 + n} \leq \frac{1}{n^3}, \quad \text{for } n = 1, 2, \dots$$

Since $\sum_{n=1}^{\infty} (1/n^3)$ is convergent, we deduce from the Comparison Test that

$$\sum_{n=1}^{\infty} \frac{1}{n^3 + n} \text{ is convergent.}$$

(b) Let

$$a_n = \frac{1}{n + \sqrt{n}}, \quad \text{for } n = 1, 2, \dots$$

Unit D3 Series

We guess that the terms of this series behave like $1/n$ for large n , and we know that $\sum_{n=1}^{\infty}(1/n)$ is divergent. We cannot compare a_n and $1/n$ directly, so we use the Limit Comparison Test with

$$b_n = \frac{1}{n}, \quad \text{for } n = 1, 2, \dots$$

Both a_n and b_n are positive and

$$\begin{aligned} \frac{a_n}{b_n} &= \frac{1}{n + \sqrt{n}} \times \frac{n}{1} \\ &= \frac{n}{n + \sqrt{n}} \\ &= \frac{1}{1 + 1/\sqrt{n}} \rightarrow 1 \neq 0. \end{aligned}$$

Since $\sum_{n=1}^{\infty}(1/n)$ is divergent, we deduce by the Limit Comparison Test that

$$\sum_{n=1}^{\infty} \frac{1}{n + \sqrt{n}} \text{ is divergent.}$$

(c) Let

$$a_n = \frac{n+4}{2n^3 - n + 1}, \quad \text{for } n = 1, 2, \dots$$

We guess that the terms of this series behave like $n/n^3 = 1/n^2$ for large n , and we know that $\sum_{n=1}^{\infty}(1/n^2)$ is convergent. We cannot compare a_n and $1/n^2$ directly, so we use the Limit Comparison Test with

$$b_n = \frac{1}{n^2}, \quad \text{for } n = 1, 2, \dots$$

Both a_n and b_n are positive and

$$\begin{aligned} \frac{a_n}{b_n} &= \frac{n+4}{2n^3 - n + 1} \times \frac{n^2}{1} \\ &= \frac{n^3 + 4n^2}{2n^3 - n + 1} \\ &= \frac{1 + 4/n}{2 - 1/n^2 + 1/n^3} \rightarrow \frac{1}{2} \neq 0. \end{aligned}$$

Since $\sum_{n=1}^{\infty}(1/n^2)$ is convergent, we deduce by the Limit Comparison Test that

$$\sum_{n=1}^{\infty} \frac{n+4}{2n^3 - n + 1} \text{ is convergent.}$$

(d) We guess that this series is dominated by $\sum_{n=1}^{\infty}(1/n^3)$. Indeed, since

$$0 \leq \cos^2(2n) \leq 1, \quad \text{for } n = 1, 2, \dots,$$

we have

$$0 \leq \frac{\cos^2(2n)}{n^3} \leq \frac{1}{n^3}, \quad \text{for } n = 1, 2, \dots$$

Since $\sum_{n=1}^{\infty}(1/n^3)$ is convergent, we deduce by the Comparison Test that

$$\sum_{n=1}^{\infty} \frac{\cos^2(2n)}{n^3} \text{ is convergent.}$$

Solution to Exercise D48

(a) Let

$$a_n = \frac{n^3}{n!}, \quad n = 1, 2, \dots$$

Then a_n is positive and

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{(n+1)^3}{(n+1)!} \times \frac{n!}{n^3} \\ &= \frac{(n+1)^3}{(n+1)n^3} \\ &= \frac{n^2 + 2n + 1}{n^3} = \frac{1}{n} + \frac{2}{n^2} + \frac{1}{n^3}. \end{aligned}$$

Hence, by the Combination Rules for sequences,

$$\frac{a_{n+1}}{a_n} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Thus, by the Ratio Test,

$$\sum_{n=1}^{\infty} \frac{n^3}{n!} \text{ is convergent.}$$

(b) Let

$$a_n = \frac{n^2 2^n}{n!}, \quad n = 1, 2, \dots$$

Then a_n is positive and

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{(n+1)^2 2^{n+1}}{(n+1)!} \times \frac{n!}{n^2 2^n} \\ &= \frac{2(n+1)^2}{(n+1)n^2} = 2 \left(\frac{1}{n} + \frac{1}{n^2} \right). \end{aligned}$$

Hence, by the Combination Rules for sequences,

$$\frac{a_{n+1}}{a_n} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Thus, by the Ratio Test,

$$\sum_{n=1}^{\infty} \frac{n^2 2^n}{n!} \text{ is convergent.}$$

(c) Let

$$a_n = \frac{(2n)!}{n^n}, \quad n = 1, 2, \dots$$

Then a_n is positive and

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{(2(n+1))!}{(n+1)^{n+1}} \times \frac{n^n}{(2n)!} \\ &= \frac{(2n+2)! n^n}{(n+1)^{n+1} (2n)!} \\ &= \frac{(2n+2)(2n+1)n^n}{(n+1)^{n+1}} \\ &= \frac{2(2n+1)n^n}{(n+1)^n} = \frac{4n+2}{(1+1/n)^n}. \end{aligned}$$

Now

$$\lim_{n \rightarrow \infty} (1 + 1/n)^n = e \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{1}{4n+2} = 0,$$

so $((1 + 1/n)^n / (4n+2))$ is null, by the Product Rule.

We deduce by the Reciprocal Rule that

$$\frac{a_{n+1}}{a_n} \rightarrow \infty \quad \text{as } n \rightarrow \infty;$$

see Subsection 4.3 of Unit D2.

Hence, by the Ratio Test,

$$\sum_{n=1}^{\infty} \frac{(2n)!}{n^n} \text{ is divergent.}$$

(Alternatively, note that

$$\begin{aligned} \frac{(2n)!}{n^n} &\geq \left(\frac{2n}{n}\right) \left(\frac{2n-1}{n}\right) \cdots \left(\frac{n+1}{n}\right) \\ &\geq 1, \end{aligned}$$

so, by the Non-null Test,

$$\sum_{n=1}^{\infty} \frac{(2n)!}{n^n} \text{ is divergent.)}$$

Solution to Exercise D49

(a) $\sum_{n=1}^{\infty} \frac{2^n}{n!}$ is a basic convergent series of type (d), with $c = 2$.

(b) $\sum_{n=1}^{\infty} \frac{1}{n^{5/2}}$ is a basic convergent series of type (a), with $p = \frac{5}{2}$.

(c) $\sum_{n=1}^{\infty} \frac{1}{n^{2/3}}$ is a basic divergent series of type (e), with $p = \frac{2}{3}$.

(d) $\sum_{n=1}^{\infty} \frac{n}{2^n}$ is a basic convergent series of type (c), with $p = 1$ and $c = \frac{1}{2}$.

(e) $\sum_{n=1}^{\infty} \frac{1}{4^n}$ is a basic convergent series of type (b), with $c = \frac{1}{4}$.

Solution to Exercise D50

(a) Let

$$a_n = \frac{(-1)^{n+1} n}{n^3 + 1}, \quad n = 1, 2, \dots$$

Then

$$|a_n| = \frac{n}{n^3 + 1}, \quad \text{for } n = 1, 2, \dots$$

Now

$$\frac{n}{n^3 + 1} \leq \frac{n}{n^3} = \frac{1}{n^2}, \quad \text{for } n = 1, 2, \dots,$$

and $\sum_{n=1}^{\infty} (1/n^2)$ is a basic convergent series. Hence, by the Comparison Test,

$$\sum_{n=1}^{\infty} \frac{n}{n^3 + 1} \text{ is convergent.}$$

By the Absolute Convergence Test, it follows that

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1} n}{n^3 + 1} \text{ is convergent.}$$

(b) If we write this series as $\sum_{n=0}^{\infty} a_n$, then

$$|a_n| = \frac{1}{2^n}, \quad \text{for } n = 0, 1, 2, \dots$$

Since $\sum_{n=0}^{\infty} (1/2^n)$ is a basic convergent series, it follows from the Absolute Convergence Test that

$$\sum_{n=0}^{\infty} a_n \text{ is convergent.}$$

Solution to Exercise D51

If we write the series as $\sum_{n=1}^{\infty} a_n$, then

$$|a_n| = \frac{1}{2^n}, \text{ for } n = 1, 2, \dots$$

Since $\sum_{n=1}^{\infty} (1/2^n)$ is a basic convergent (geometric) series, it follows from the Absolute Convergence Test that $\sum_{n=1}^{\infty} a_n$ is convergent.

By the infinite form of the Triangle Inequality, we have

$$\left| \sum_{n=1}^{\infty} a_n \right| \leq \sum_{n=1}^{\infty} |a_n| = \sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1/2}{1 - 1/2} = 1.$$

Hence the sum lies in the interval $[-1, 1]$.

Solution to Exercise D52

(a) Let

$$a_n = \frac{(-1)^{n+1}}{n^{1/3}}, \text{ for } n = 1, 2, \dots$$

Then $a_n = (-1)^{n+1}b_n$, where

$$b_n = \frac{1}{n^{1/3}}, \text{ for } n = 1, 2, \dots$$

Now

1. $b_n \geq 0$, for $n = 1, 2, \dots$
2. (b_n) is a basic null sequence
3. (b_n) is decreasing, because $(1/b_n) = (n^{1/3})$ is increasing.

Hence, by the Alternating Test,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^{1/3}} \text{ is convergent.}$$

(b) Let

$$a_n = \frac{(-1)^{n+1}}{n + n^{1/2}}, \text{ for } n = 1, 2, \dots$$

Then $a_n = (-1)^{n+1}b_n$, where

$$b_n = \frac{1}{n + n^{1/2}}, \text{ for } n = 1, 2, \dots$$

Now

1. $b_n \geq 0$, for $n = 1, 2, \dots$
2. (b_n) is a null sequence by the Squeeze Rule, since

$$0 \leq \frac{1}{n + n^{1/2}} \leq \frac{1}{n}, \text{ for } n = 1, 2, \dots$$

3. (b_n) is decreasing, because $(1/b_n) = (n + n^{1/2})$ is increasing.

Hence, by the Alternating Test,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n + n^{1/2}} \text{ is convergent.}$$

(c) Let

$$a_n = \frac{(-1)^{n+1}n}{n + 2}, \text{ for } n = 1, 2, \dots$$

Then

$$|a_n| = \frac{n}{n + 2} = \frac{1}{1 + 2/n} \rightarrow 1 \neq 0.$$

Hence, by the Non-null Test,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}n}{n + 2} \text{ is divergent.}$$

(Notice that if you try to apply the Alternating Test here, by writing $a_n = (-1)^{n+1}b_n$, where

$$b_n = \frac{n}{n + 2}, \text{ for } n = 1, 2, \dots,$$

then you find that (b_n) is not null, so the Alternating Test cannot be used. Indeed, the series $\sum_{n=1}^{\infty} a_n$ is divergent, as the Non-null Test shows.)

Solution to Exercise D53

(a) We have

$$\frac{5n + 2^n}{3^n} = 5n\left(\frac{1}{3}\right)^n + \left(\frac{2}{3}\right)^n, \text{ for } n = 1, 2, \dots$$

Now

$$\sum_{n=1}^{\infty} n\left(\frac{1}{3}\right)^n \text{ and } \sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^n$$

are both basic convergent series. Thus, by the Combination Rules for series,

$$\sum_{n=1}^{\infty} \frac{5n + 2^n}{3^n} \text{ is convergent.}$$

(b) We guess that the terms of the series behave like $1/n^3$ for large n , so we use the Limit Comparison Test with

$$a_n = \frac{3}{2n^3 - 1}, \text{ for } n = 1, 2, \dots$$

and

$$b_n = 1/n^3, \text{ for } n = 1, 2, \dots$$

Both a_n and b_n are positive and

$$\begin{aligned}\frac{a_n}{b_n} &= \frac{3}{2n^3 - 1} \times \frac{n^3}{1} \\ &= \frac{3n^3}{2n^3 - 1} = \frac{3}{2 - 1/n^3} \rightarrow \frac{3}{2} \neq 0.\end{aligned}$$

Since $\sum_{n=1}^{\infty} (1/n^3)$ is a basic convergent series, we deduce from the Limit Comparison Test that

$$\sum_{n=1}^{\infty} \frac{3}{2n^3 - 1} \text{ is convergent.}$$

(c) We use the Alternating Test.

Let

$$a_n = \frac{(-1)^{n+1}}{n \log(n+1)}, \quad \text{for } n = 1, 2, \dots$$

Then $a_n = (-1)^{n+1} b_n$, where

$$b_n = \frac{1}{n \log(n+1)}, \quad \text{for } n = 1, 2, \dots$$

Now

1. $b_n \geq 0$, for $n = 1, 2, \dots$
 2. (b_n) is a null sequence, by the Squeeze Rule, since
- $$0 \leq \frac{1}{n \log(n+1)} \leq \frac{1}{n \log 2}, \quad \text{for } n = 1, 2, \dots$$
3. (b_n) is decreasing, because $(1/b_n) = (n \log(n+1))$ is increasing.

Hence, by the Alternating Test,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n \log(n+1)} \text{ is convergent.}$$

(d) Let

$$a_n = \frac{(-1)^{n+1} n^2}{n^2 + 1}, \quad \text{for } n = 1, 2, \dots$$

Then

$$|a_n| = \frac{n^2}{n^2 + 1} = \frac{1}{1 + 1/n^2} \rightarrow 1 \neq 0.$$

Hence, by the Non-null Test,

$$\sum_{n=1}^{\infty} a_n \text{ is divergent.}$$

(e) Let

$$a_n = \frac{(-1)^{n+1} n}{n^3 + 5}, \quad \text{for } n = 1, 2, \dots$$

Then

$$|a_n| = \frac{n}{n^3 + 5}, \quad \text{for } n = 1, 2, \dots$$

Thus

$$|a_n| \leq \frac{n}{n^3} = \frac{1}{n^2}, \quad \text{for } n = 1, 2, \dots$$

Since $\sum(1/n^2)$ is a basic convergent series, we deduce by the Comparison Test that

$$\sum_{n=1}^{\infty} |a_n| \text{ is convergent.}$$

Hence, by the Absolute Convergence Test,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1} n}{n^3 + 5} \text{ is convergent.}$$

(f) Let

$$a_n = \frac{2^n}{n^6}, \quad \text{for } n = 1, 2, \dots$$

Then a_n is positive and

$$\frac{1}{a_n} = \frac{n^6}{2^n} \rightarrow 0,$$

since $(n^6/2^n)$ is a basic null sequence. So, by the Reciprocal Rule, $a_n \rightarrow \infty$. Hence, by the Non-null Test,

$$\sum_{n=1}^{\infty} \frac{2^n}{n^6} \text{ is divergent.}$$

Unit D4

Continuity

Introduction

In Unit A4 *Real functions, graphs and conics* you studied techniques for sketching the graphs of many common real functions. For this purpose, we made various assumptions about these graphs – in particular, that they have no gaps (except ‘obvious’ ones at asymptotes). In this unit you will see how to justify this assumption by using the concept of a *continuous function* and showing that many familiar functions are continuous. This concept is important in analysis because, in many cases, the easiest way to prove that a function has a property which may seem intuitively obvious is to use the fact that the function is continuous.

1 Operations on functions

This section gives an overview of the various fundamental operations on functions: forming combinations, composites and inverses of functions. These operations were discussed in more detail in Book A *Introduction*.

We begin with a brief review of notation and basic terminology. In this unit we are concerned with *real functions*, that is, functions whose domain and codomain are subsets of \mathbb{R} . Such functions can be specified in various ways. For example, the function

$$\begin{aligned} f : \mathbb{R} - \{0\} &\longrightarrow \mathbb{R} \\ x &\longmapsto 1/x \end{aligned}$$

can also be written as

$$f(x) = 1/x \quad (x \in \mathbb{R} - \{0\}),$$

where the codomain of f is assumed to be \mathbb{R} . It can also be written simply as

$$f(x) = 1/x,$$

where now the domain of f is assumed to be the set of values of x for which $1/x$ is defined, that is, $\mathbb{R} - \{0\}$, and where the codomain is \mathbb{R} . These notations illustrate the following convention, which you met in Unit A4.

Convention for real functions

When a real function is specified *only by a rule*, it is to be understood that the domain of the function is the set of all real numbers for which the rule is applicable, and the codomain of the function is \mathbb{R} .

Recall that, if A and B are subsets of \mathbb{R} and $f : A \rightarrow B$ is a real function, then

- the **image set** of f is the set $f(A) = \{f(x) : x \in A\}$
- f is **one-to-one** if each element of $f(A)$ is the image of exactly one element of A
- f is **onto** if $f(A) = B$.

The above convention for real functions is concise and we often use it. Sometimes, however, if we want to assert that a function has a particular property, then we may need to restrict its domain or codomain. For example, the function

$$f(x) = \sin x$$

has domain and codomain \mathbb{R} , by our convention, but it is neither one-to-one nor onto. However, the function

$$\begin{aligned} g : [-\pi/2, \pi/2] &\longrightarrow [-1, 1] \\ x &\longmapsto \sin x \end{aligned}$$

has the same rule as the above function f , but g is both one-to-one and onto.

When we say that a function f is *defined on* a set I (usually an interval), this means that the domain of f contains the set I . For example, the function $f(x) = 1/x$ is defined on $[1, 2]$, but not on $[-1, 1]$. The definitions and notation for the various types of interval (open, closed, half-open) are given in the module Handbook; we make frequent use of them in this unit.

1.1 Sums, products and quotients of functions

Let f and g be the functions

$$f(x) = 1/x \quad (x \in \mathbb{R} - \{0\}) \quad \text{and} \quad g(x) = \sin x \quad (x \in \mathbb{R}).$$

The graphs of these functions are shown in Figure 1.

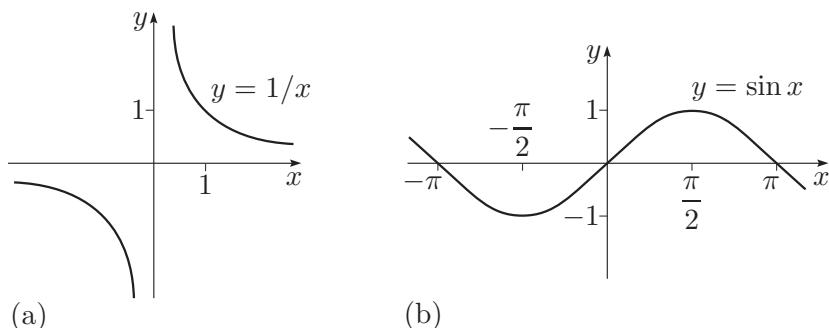


Figure 1 The graphs of (a) $y = 1/x$ and (b) $y = \sin x$

We use the expressions $f + g$, fg and f/g to denote the functions

$$(f + g)(x) = f(x) + g(x) = 1/x + \sin x \quad (x \in \mathbb{R} - \{0\}),$$

$$(fg)(x) = f(x)g(x) = \frac{\sin x}{x} \quad (x \in \mathbb{R} - \{0\})$$

and

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)} = \frac{1}{x \sin x} \quad (x \in \mathbb{R} - \{n\pi : n \in \mathbb{Z}\}).$$

The domains of $f + g$, fg and f/g include only those points at which f and g are *both* defined. (We often use the word ‘point’ to mean ‘number’.) Also, when forming the quotient f/g , we must exclude from the domain all the points x such that $g(x) = 0$. The formal definitions are as follows.

Definitions

Let A and B be subsets of \mathbb{R} and $f : A \rightarrow \mathbb{R}$ and $g : B \rightarrow \mathbb{R}$ be functions. Then

- the **sum** $f + g$ is the function with domain $A \cap B$ and rule

$$(f + g)(x) = f(x) + g(x)$$

- for $\lambda \in \mathbb{R}$, the **multiple** λf is the function with domain A and rule

$$(\lambda f)(x) = \lambda f(x)$$

- the **product** fg is the function with domain $A \cap B$ and rule

$$(fg)(x) = f(x)g(x)$$

- the **quotient** f/g is the function with domain

$$A \cap B - \{x : g(x) = 0\}$$

and rule

$$(f/g)(x) = f(x)/g(x).$$

Often we wish to form the sum, product or quotient of functions f and g which have the *same* domain, A say. In this case, A is also the domain of $f + g$ and fg , and the domain of f/g is $A - \{x : g(x) = 0\}$.

Exercise D54

Let f and g be the functions

$$f(x) = e^x \quad (x \in \mathbb{R}) \quad \text{and} \quad g(x) = \tan x \quad (x \in (-\pi/2, \pi/2)).$$

Determine the domain and rule of the functions $f + g$, fg and f/g .

1.2 Composite functions

Let f and g be functions. Then the composite function $g \circ f$ is the function defined by the rule

$$(g \circ f)(x) = g(f(x)),$$

where we apply first f and then g . Again, we must exclude from the domain all the points x which lead to an expression that is not defined.

For example, if

$$f(x) = 1/x \quad (x \in \mathbb{R} - \{0\}) \quad \text{and} \quad g(x) = \sin x \quad (x \in \mathbb{R}),$$

then $g \circ f$ is the function

$$(g \circ f)(x) = \sin \frac{1}{x} \quad (x \in \mathbb{R} - \{0\}),$$

whereas $f \circ g$ is the function

$$(f \circ g)(x) = \frac{1}{\sin x} = \operatorname{cosec} x \quad (x \in \mathbb{R} - \{n\pi : n \in \mathbb{Z}\}).$$

The graphs of $g \circ f$ and $f \circ g$ are shown in Figure 2.

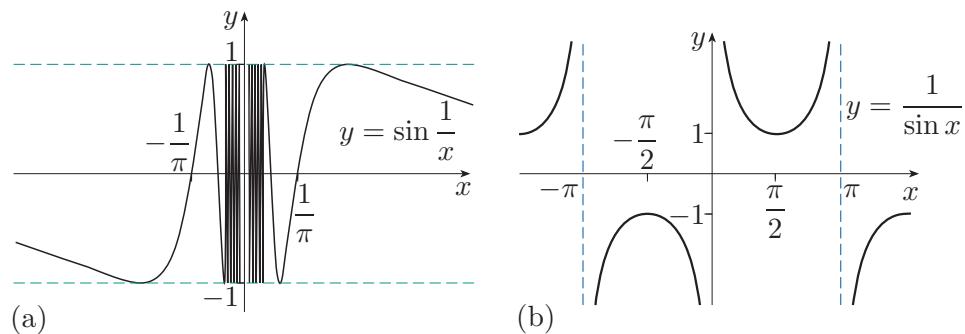


Figure 2 The graphs of (a) $y = \sin \frac{1}{x}$ and (b) $y = \frac{1}{\sin x}$

More generally, if $f : A \rightarrow \mathbb{R}$ and $g : B \rightarrow \mathbb{R}$, then $g(f(x))$ is defined if and only if x lies in the domain of f and $f(x)$ lies in the domain of g , as illustrated in Figure 3.

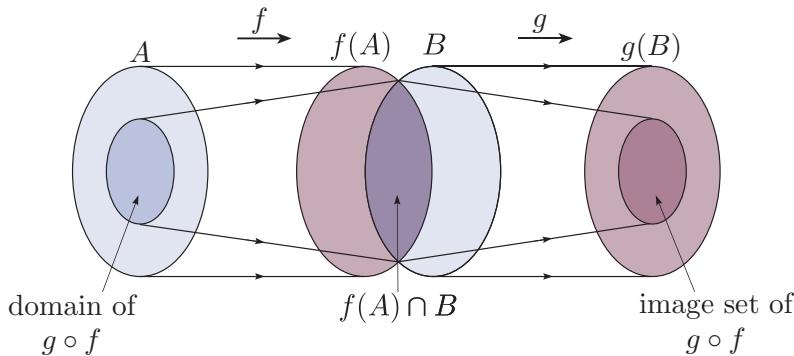


Figure 3 The domain and image set of a composite function

Definition

Let A and B be subsets of \mathbb{R} and $f : A \rightarrow \mathbb{R}$ and $g : B \rightarrow \mathbb{R}$ be functions. Then the **composite** $g \circ f$ has domain

$$\{x \in A : f(x) \in B\}$$

and rule

$$(g \circ f)(x) = g(f(x)).$$

This definition allows us to form the composite of *any* two functions, though in some cases the domain of the composite is the empty set \emptyset . For example, if

$$f(x) = -x^2 - 1 \quad (x \in \mathbb{R}) \quad \text{and} \quad g(x) = \sqrt{x} \quad (x \in [0, \infty)),$$

then $f(\mathbb{R}) \subset (-\infty, -1]$ which has empty intersection with $[0, \infty)$, the domain of g . So the domain of the composite function $g \circ f$ is \emptyset .

Frequently, however, it happens that the image set of f is a subset of the domain of g ; that is, $f(A) \subseteq B$. In this case, the set A is also the domain of $g \circ f$.

Exercise D55

Let f and g be the functions

$$f(x) = \sqrt{x} \quad (x \in [0, \infty)) \quad \text{and} \quad g(x) = \sin x \quad (x \in \mathbb{R}).$$

Determine the domain and rule of the composites $f \circ g$ and $g \circ f$.

1.3 Inverse functions

Let f be the function

$$f(x) = 2x \quad (x \in \mathbb{R}).$$

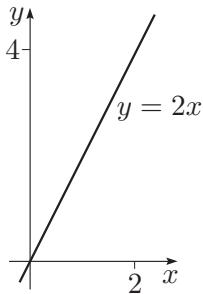


Figure 4 The graph of $y = 2x$

The graph of f is shown in Figure 4. For each number y in \mathbb{R} , there is a unique number $x = y/2$ in the domain of f such that

$$f(x) = f\left(\frac{y}{2}\right) = 2 \times \frac{y}{2} = y.$$

The corresponding function $g(y) = y/2$ is called the *inverse function* of f because it undoes the effect of f ; that is,

$$g(f(x)) = x, \quad \text{for } x \in \mathbb{R},$$

and

$$f(g(y)) = y, \quad \text{for } y \in \mathbb{R}.$$

However, not every function has an inverse function. For example, consider the function

$$f(x) = x^2 \quad (x \in \mathbb{R}).$$

Since $f(2) = 4 = f(-2)$, we cannot assign a unique value x in the domain of f such that $f(x) = 4$. The problem here is that this function f is not one-to-one. In general, it is possible to define the inverse function of a function only if that function is one-to-one.

We now give the definition of the inverse function, illustrated in Figure 5.

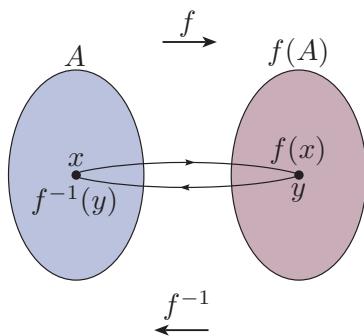


Figure 5 The inverse function

For some functions f , we can find the inverse function f^{-1} directly by solving the equation $y = f(x)$ algebraically to obtain x in terms of y .

Worked Exercise D40

Prove that the following function has an inverse function, and find the domain and rule of this inverse function:

$$f(x) = \frac{1}{1-x} \quad (x \in (-\infty, 1)).$$

Solution

We solve the equation $y = f(x)$ to obtain x in terms of y .

Let

$$y = \frac{1}{1-x}.$$

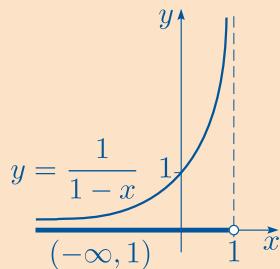
Rearranging this equation, we obtain

$$\begin{aligned} y = \frac{1}{1-x} &\iff \frac{1}{y} = 1-x \\ &\iff x = 1 - \frac{1}{y}. \end{aligned}$$

This means that each value of y is the image of exactly one value of x , namely $x = 1 - 1/y$.

This shows that f is one-to-one, so f has an inverse function with rule $f^{-1}(y) = 1 - 1/y$.

To determine the domain of f^{-1} we must find the image set of f . A sketch of the graph of f is shown below.



From the graph, it appears that the image set of f is $(0, \infty)$. To prove this, we first show that $f((-\infty, 1))$ is a subset of $(0, \infty)$.

For each x in the domain $(-\infty, 1)$, we have $x < 1$, so

$$f(x) = \frac{1}{1-x} > 0.$$

Thus $f((-\infty, 1)) \subseteq (0, \infty)$.

Next we show that $(0, \infty)$ is a subset of $f((-\infty, 1))$. Taken together, these results show that the image set of f is $(0, \infty)$.

On the other hand, for each y in $(0, \infty)$, we have $1/y > 0$, so

$$x = 1 - \frac{1}{y} \in (-\infty, 1).$$

Thus $f((-\infty, 1)) \supseteq (0, \infty)$, so it follows that $f((-\infty, 1)) = (0, \infty)$.

Hence the domain of f^{-1} is $(0, \infty)$, so

$$f^{-1}(x) = 1 - \frac{1}{x} \quad (x \in (0, \infty)).$$

We have used x instead of y here to conform with the usual practice of writing x for the domain variable when defining a function.

Notice that the graph $y = f^{-1}(x)$ is always obtained by reflecting the graph $y = f(x)$ in the line $y = x$. This is illustrated in Figure 6 for the functions f and f^{-1} from Worked Exercise D40.

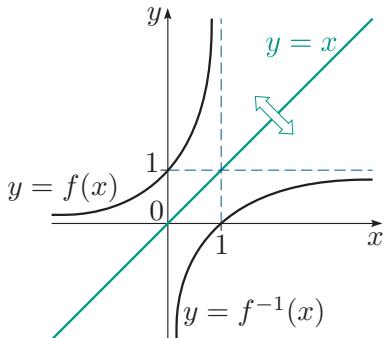


Figure 6 The graphs of the functions in Worked Exercise D40

Exercise D56

Prove that the following function has an inverse function, and find the domain and rule of this inverse function.

$$f(x) = \frac{x+3}{x-2} \quad (x \in (2, \infty))$$

Hint: It may help to write

$$\frac{x+3}{x-2} = 1 + \frac{5}{x-2}.$$

Proving that a function is one-to-one

We have seen that if $f : A \rightarrow \mathbb{R}$ is one-to-one, then f has an inverse function f^{-1} with domain $f(A)$. For the function f considered in Worked Exercise D40, it was possible to determine f^{-1} explicitly by solving the equation $y = f(x)$ to obtain x in terms of y . Unfortunately, it is often impossible to solve the equation $y = f(x)$ in this way.

Nevertheless, it may still be possible to prove that f has an inverse function f^{-1} by showing that f is one-to-one by some other method. For example in Unit A1 *Sets, functions and vectors* we showed that various functions f are one-to-one by proving algebraically that

$$\text{if } f(x_1) = f(x_2), \text{ then } x_1 = x_2.$$

However, this algebraic method can only be used for fairly simple functions.

Another way of showing that f is one-to-one is by proving that f is *strictly increasing* or *strictly decreasing*. You met these concepts applied to real functions in Unit A4 and to sequences in Unit D2, where you also encountered the idea of a *monotonic* sequence. It is now helpful to use the term monotonic in the case of real functions also. The formal definitions are given in the box below, and are illustrated in the cases of strictly increasing and strictly decreasing functions in Figures 7 and 8 respectively.

Definitions

A real function f defined on an interval I is

- **increasing** on I if

$$x_1 < x_2 \implies f(x_1) \leq f(x_2), \quad \text{for } x_1, x_2 \in I$$

- **strictly increasing** on I if

$$x_1 < x_2 \implies f(x_1) < f(x_2), \quad \text{for } x_1, x_2 \in I$$

- **decreasing** on I if

$$x_1 < x_2 \implies f(x_1) \geq f(x_2), \quad \text{for } x_1, x_2 \in I$$

- **strictly decreasing** on I if

$$x_1 < x_2 \implies f(x_1) > f(x_2), \quad \text{for } x_1, x_2 \in I$$

- **monotonic** on I if f is either increasing on I or decreasing on I

- **strictly monotonic** on I if f is either strictly increasing on I or strictly decreasing on I .

If the interval I in the definitions is the domain of f , then we omit ‘on I ’ and just say, for example,

f is increasing.

The most powerful technique for proving that a function f is increasing or decreasing is to compute the derivative f' of f and examine the sign of $f'(x)$. We used this technique in Unit A4 and we will use it again in Book F *Analysis 2*, once we have laid a rigorous foundation for the idea of the derivative of a function. For the present, however, we consider only those functions which can be proved to be increasing or decreasing by manipulating inequalities using the rules from Unit D1 *Numbers*, rather than by using calculus.

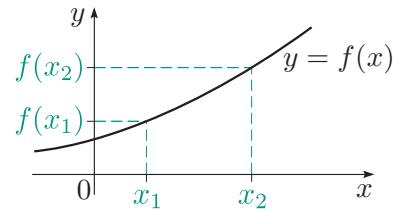


Figure 7 f is strictly increasing

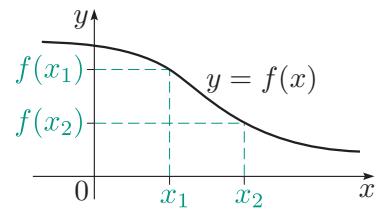


Figure 8 f is strictly decreasing

For example, if $n \in \mathbb{N}$, then the function

$$f(x) = x^n \quad (x \in [0, \infty))$$

is strictly increasing; and if n is odd, then the function

$$f(x) = x^n \quad (x \in \mathbb{R})$$

is strictly increasing; see Figure 9(a).

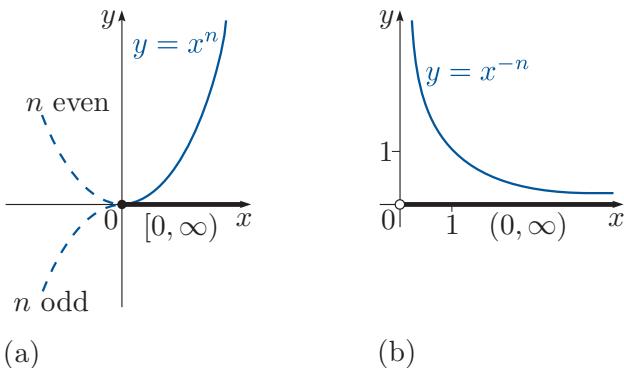


Figure 9 The graphs of (a) $y = x^n$ and (b) $y = x^{-n}$

Similarly, if $n \in \mathbb{N}$, then the function

$$f(x) = x^{-n} \quad (x \in (0, \infty))$$

is strictly decreasing; see Figure 9(b).

Any function that is strictly monotonic must be one-to-one since, if $x_1 < x_2$, then it is impossible to have $f(x_1) = f(x_2)$. Here is an example of how this property can be used.

Worked Exercise D41

Prove that the function

$$f(x) = x^5 + x - 1 \quad (x \in \mathbb{R})$$

is one-to-one.

Solution

If $x_1 < x_2$, then $x_1^5 < x_2^5$, so

$$x_1^5 + x_1 - 1 < x_2^5 + x_2 - 1; \quad \text{that is, } f(x_1) < f(x_2).$$

Hence f is strictly increasing and thus one-to-one.

Exercise D57

Prove that the following functions are one-to-one.

$$(a) \quad f(x) = x^4 + 2x + 3 \quad (x \in [0, \infty))$$

$$(b) \quad f(x) = \frac{1}{x} - x^2 \quad (x \in (0, \infty))$$

Determining the image set

If the function $f : A \rightarrow \mathbb{R}$ is strictly increasing or strictly decreasing, then f is one-to-one and so has an inverse function f^{-1} with domain $f(A)$.

However, it is not always easy to determine the image set $f(A)$.

For example, consider the function

$$f(x) = x^5 + x - 1 \quad (x \in \mathbb{R}).$$

We saw in Worked Exercise D41 that f is one-to-one, so f has an inverse function with domain $f(\mathbb{R})$. Since f is strictly increasing, it seems likely that $f : \mathbb{R} \rightarrow \mathbb{R}$ is onto, so $f(\mathbb{R}) = \mathbb{R}$, and we have sketched the graph $y = f(x)$ in Figure 10 as though this is the case. But how can we prove that $f(\mathbb{R}) = \mathbb{R}$? To do this we want to show that, for each $y \in \mathbb{R}$, there is an x such that

$$f(x) = x^5 + x - 1 = y.$$

Unfortunately, we cannot find such an x by solving this equation algebraically to obtain x in terms of y . Could it be that the graph $y = x^5 + x - 1$ actually has some ‘gaps’ or ‘jumps’ in it? We would be very surprised if gaps do occur, but how can we prove that they do not?

To answer this question, we need the concept of *continuity*, which we introduce in Section 2.

2 Continuous functions

In this section you will see what it means for a function f to be continuous at a point a . You will also meet several rules which enable you to combine continuous functions in various ways to form other continuous functions.

Using these rules, together with a list of basic continuous functions, we can deduce that many functions are continuous at each point of their domains.

For example, the function

$$x \mapsto x \sin(1/x)$$

is continuous at each point of $\mathbb{R} - \{0\}$.

You will also meet rules which enable us to prove that certain hybrid functions are continuous. For example, we can show that the function

$$f(x) = \begin{cases} x \sin(1/x), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is continuous at 0. The graph of this function is shown in Figure 11.

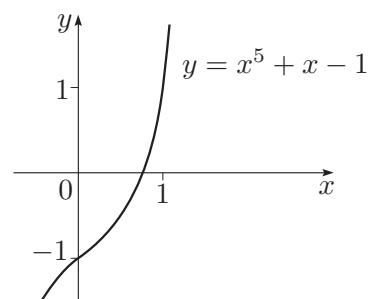


Figure 10 The graph of $y = x^5 + x - 1$

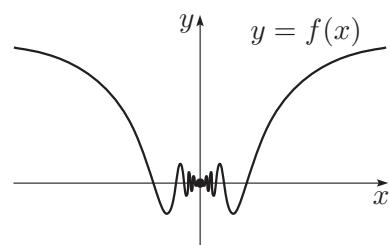


Figure 11 The graph of a continuous hybrid function

2.1 What is continuity?

At the end of Section 1 we asked whether the graph of the function

$$f(x) = x^5 + x - 1 \quad (x \in \mathbb{R})$$

has any gaps. Roughly speaking, can we draw the graph $y = x^5 + x - 1$ without lifting a pen from the paper? In this section we show that this graph cannot have any gaps because the function f is *continuous*.

Our first objective is to *define* the phrase

f is continuous at the point a .

To agree with our intuitive ideas, we wish to define this concept in such a way that the two functions shown in Figure 12 are continuous at the point a .

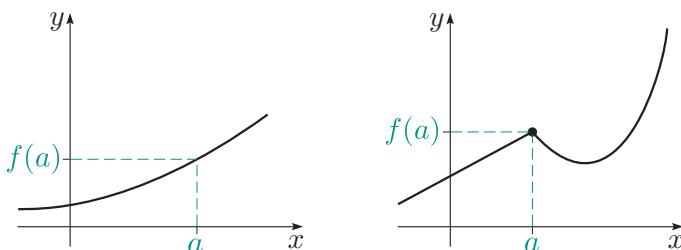


Figure 12 The graphs of two functions which are continuous at a

On the other hand, we wish to formulate our definition so that the two functions shown in Figure 13 are *not* continuous at the point a .

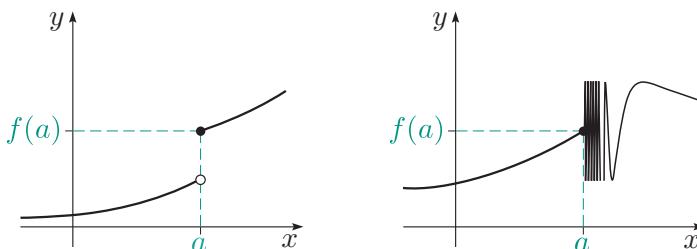


Figure 13 The graphs of two functions which are not continuous at a

For each of the graphs in Figure 12, as the values of x get closer and closer to a , the corresponding values of $f(x)$ get closer and closer to $f(a)$. This is not the case for the graphs in Figure 13. So our definition of continuity must say, in a precise way, that

if x tends to a , then $f(x)$ tends to $f(a)$.

There are several ways of making this idea precise. Here we adopt a definition that involves the *convergence of sequences*, since this enables us to use results about sequences that you have already met to prove that many common functions are continuous. (In Book F we give another definition of continuity which is convenient for dealing with more unusual functions.)

To motivate our definition of continuity we start with the following exercise. Here the sequences are denoted by (x_n) rather than (a_n) because, as you will see, they represent points on the x -axis.

Exercise D58

Let (x_n) be a sequence such that $\lim_{n \rightarrow \infty} x_n = 2$. Determine the limits of the following sequences.

- (a) $(3x_n)$ (b) (x_n^2) (c) $(1/x_n)$

We now look again at these three limits, this time from a geometrical point of view.

We saw, in part (a) of Exercise D58, that

if $x_n \rightarrow 2$, then $3x_n \rightarrow 6$.

Using function notation with $f(x) = 3x$, we can express this statement in the equivalent form

if $x_n \rightarrow 2$, then $f(x_n) \rightarrow f(2)$.

This is illustrated in Figure 14.

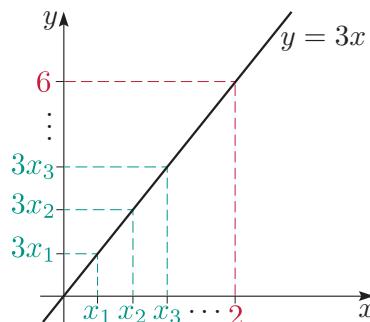


Figure 14 The graph of $y = 3x$

Next, we saw in part (b) of Exercise D58 that

$$\text{if } x_n \rightarrow 2, \text{ then } x_n^2 \rightarrow 4.$$

Using function notation with $f(x) = x^2$, we can express this statement in the equivalent form

$$\text{if } x_n \rightarrow 2, \text{ then } f(x_n) \rightarrow f(2).$$

This is illustrated in Figure 15.

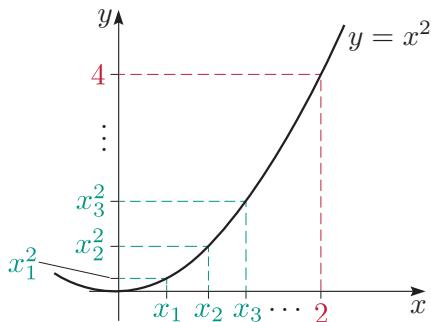


Figure 15 The graph of $y = x^2$

Finally, we saw, in part (c) of Exercise D58 that

$$\text{if } x_n \rightarrow 2, \text{ then } 1/x_n \rightarrow 1/2.$$

Using function notation with $f(x) = 1/x$, we can express this statement in the equivalent form

$$\text{if } x_n \rightarrow 2, \text{ then } f(x_n) \rightarrow f(2).$$

This is illustrated in Figure 16.

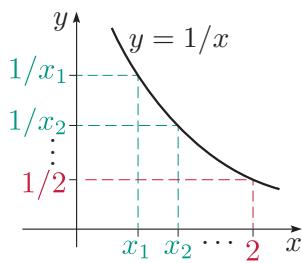


Figure 16 The graph of $y = 1/x$

So in each of these three cases, for any sequence (x_n) that has limit 2, we have shown that the sequence $f(x_n)$ has limit $f(2)$; that is,

$$\text{if } x_n \rightarrow 2, \text{ then } f(x_n) \rightarrow f(2).$$

In the above figures we have illustrated the situation in the case that the sequence (x_n) is increasing, but the conclusion holds no matter how (x_n) approaches 2.

In general, for a function $f : A \rightarrow \mathbb{R}$, where A is a subset of \mathbb{R} and $a \in A$, our formal definition of the continuity of f at a should encapsulate the

intuitive notion of continuity that, if (x_n) is any sequence such that x_n tends to a , then $f(x_n)$ tends to $f(a)$.

Our definition should also enable us to conclude that, if there is a sequence (x_n) such that x_n tends to a but $f(x_n)$ does *not* tend to $f(a)$, then f is *not continuous* at a . In such a situation we say that f is *discontinuous* at a .

For instance, it may be that the graph of f near a has a jump at a .

These two different situations are illustrated in Figure 17.

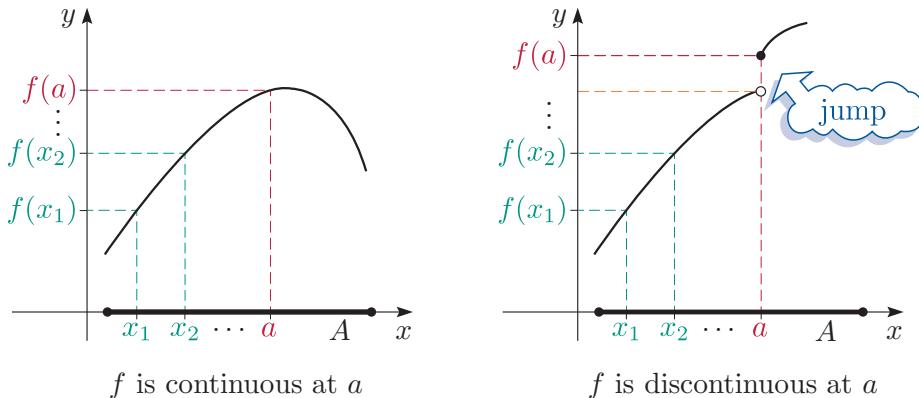


Figure 17 Continuity in terms of sequences

We now give our formal definition of continuity.

Definitions

A function $f : A \rightarrow \mathbb{R}$ is **continuous** at a point $a \in A$ if for each sequence (x_n) in A such that $x_n \rightarrow a$, we have $f(x_n) \rightarrow f(a)$.

We say that f is **continuous** (on A) if f is continuous at each point $a \in A$.

If f is not continuous at a point a in A , then we say that f is **discontinuous** at a .

We can write the above condition for continuity more concisely as follows:

$$x_n \rightarrow a \implies f(x_n) \rightarrow f(a), \text{ where } (x_n) \text{ lies in } A.$$

The next two worked exercises illustrate how we can use the definition of continuity to show that a function is continuous at a given point or discontinuous at a given point.

Worked Exercise D42

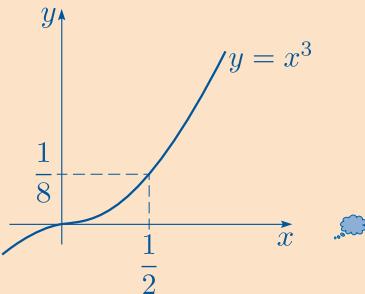
Prove that the function

$$f(x) = x^3 \quad (x \in \mathbb{R})$$

is continuous at the point $1/2$.

Solution

The graph of f is shown below. It appears from the graph that f is continuous at $1/2$ but we must now prove that this is true.



Let (x_n) be *any* sequence in \mathbb{R} that converges to $1/2$; that is, $x_n \rightarrow 1/2$ as $n \rightarrow \infty$. Then, by the Product Rule for sequences, it follows that

$$f(x_n) = x_n^3 \rightarrow (1/2)^3 = 1/8 \text{ as } n \rightarrow \infty.$$

Since $f(1/2) = 1/8$, we have

$$f(x_n) \rightarrow f(1/2) \text{ as } n \rightarrow \infty.$$

It follows that f is continuous at $1/2$, as required.

Worked Exercise D43

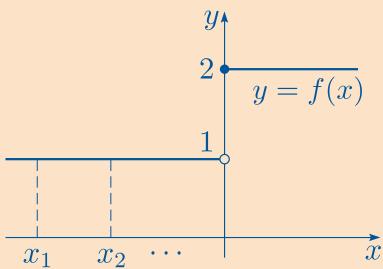
Prove that the function

$$f(x) = \begin{cases} 1, & x < 0, \\ 2, & x \geq 0, \end{cases}$$

is discontinuous at 0.

Solution

In order to prove that a function is discontinuous at a particular point, we just need to find *one* sequence for which our definition of continuity at the relevant point does *not* hold. We begin by looking at the graph of f to find a suitable sequence.



Let (x_n) be any sequence in \mathbb{R} that converges to 0 from the *right* as $n \rightarrow \infty$. By looking at the graph of f , it is clear that for such a sequence $f(x_n) \rightarrow 2 = f(0)$. This won't help us to show that f is discontinuous at 0. However, if we let (x_n) be any sequence in \mathbb{R} that converges to 0 from the *left*, the situation is very different. By looking at the graph of f , it is clear that, for such a sequence, $f(x_n) \rightarrow 1 \neq f(0)$. We choose $x_n = -1/n$ as this is a simple sequence of this type. We now make this reasoning precise. 

We first note that $f(0) = 2$. We now choose

$$x_n = -1/n, \quad n = 1, 2, \dots$$

Then $x_n \rightarrow 0$ and, since $x_n < 0$, we have $f(x_n) = 1$ for $n = 1, 2, \dots$, so

$$f(x_n) \rightarrow 1 \neq f(0) \text{ as } n \rightarrow \infty.$$

Hence f is discontinuous at 0.

Worked Exercises D42 and D43 illustrate the following general strategy.

Strategy D14

- To prove that a function $f : A \rightarrow \mathbb{R}$ is *continuous* at the point $a \in A$:
show that, if (x_n) is *any* sequence in A such that $x_n \rightarrow a$, then $f(x_n) \rightarrow f(a)$.
- To prove that a function $f : A \rightarrow \mathbb{R}$ is *discontinuous* at the point $a \in A$:
find *one* sequence (x_n) in A such that $x_n \rightarrow a$ but $f(x_n) \not\rightarrow f(a)$.

Remarks

1. The symbol $\not\rightarrow$ is read as 'does not tend to'.
2. When proving discontinuity at a point a it is often possible to choose the sequence to be $(a - 1/n)$ or $(a + 1/n)$.

We now illustrate the use of this strategy.

Worked Exercise D44

- (a) Determine whether the function

$$f(x) = 1/x \quad (x \in \mathbb{R} - \{0\})$$

is continuous.

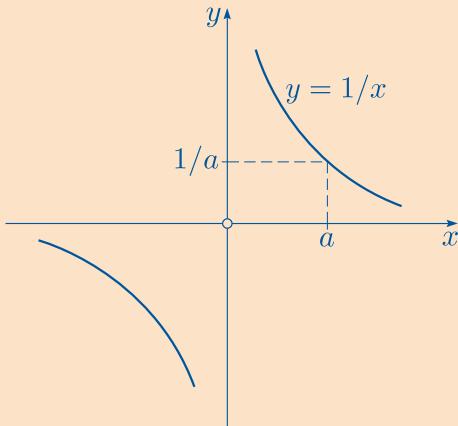
- (b) Determine whether the function

$$f(x) = \begin{cases} 1/x, & x \neq 0, \\ 0, & x = 0 \end{cases}$$

is continuous at 0.

Solution

- (a) Recall that a function is continuous if it is continuous at each point in its domain. The graph suggests that f is continuous everywhere it is defined.



We now prove that this is the case using Strategy D14.

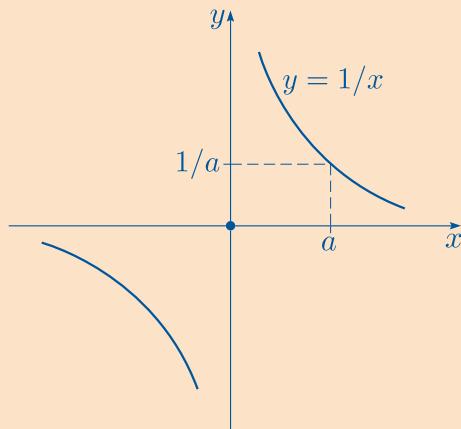
The function f has domain $\mathbb{R} - \{0\}$.

If $a \neq 0$ and (x_n) is a sequence in $\mathbb{R} - \{0\}$ with $x_n \rightarrow a$, then, by the Quotient Rule for sequences,

$$f(x_n) = 1/x_n \rightarrow 1/a = f(a).$$

It follows that f is continuous at every $a \neq 0$ and hence at every point in its domain.

- (b)  This function is the same as in part (a) apart from at 0 which is now in the domain. The graph of f has a jump at 0, so it certainly looks as though f is discontinuous at 0.



We now prove this using Strategy D14. 

We first note that $f(0) = 0$.

We now choose

$$x_n = 1/n, \quad n = 1, 2, \dots$$

Then $x_n \rightarrow 0$ but

$$f(x_n) = f(1/n) = n \rightarrow \infty,$$

so $f(x_n) \not\rightarrow f(0) = 0$. It follows that the function f is discontinuous at 0.

Exercise D59

Determine whether the following functions are continuous at the points given. (Remember that $\lfloor x \rfloor$ denotes the integer part of x .)

- (a) $f(x) = x^3 - 2x^2$, at the point $a = 2$
 (b) $f(x) = \lfloor x \rfloor$, at the point $a = 1$

Exercise D60

Prove that the following functions are continuous (that is, continuous at every point $a \in \mathbb{R}$).

- (a) $f(x) = 1$
 (b) $f(x) = x$

We conclude this subsection with two worked exercises in which we prove the continuity of two important functions: the modulus function and the square root function.

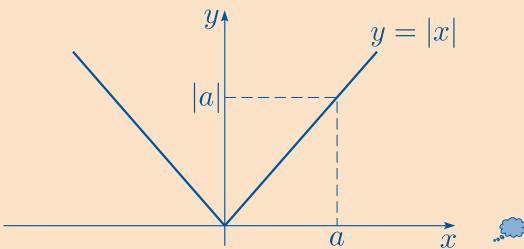
Worked Exercise D45

Determine whether the following function is continuous.

$$f(x) = |x| \quad (x \in \mathbb{R})$$

Solution

It certainly seems from the graph that f is continuous on its domain \mathbb{R} .



We guess that f is continuous on its domain \mathbb{R} . Let $a \in \mathbb{R}$, and let (x_n) be any sequence in \mathbb{R} with $x_n \rightarrow a$ as $n \rightarrow \infty$. We want to prove that

$$x_n \rightarrow a \implies |x_n| \rightarrow |a|,$$

or, in other words, that

$$(x_n - a) \text{ is a null sequence} \implies (|x_n| - |a|) \text{ is a null sequence. (*)}$$

To prove statement (*), we use the backwards form of the Triangle Inequality that you met in Unit D1. This says that for any $a, b \in \mathbb{R}$, we have $|a - b| \geq ||a| - |b||$.

Now it follows from the backwards form of the Triangle Inequality that

$$|x_n - a| \geq ||x_n| - |a||, \quad \text{for } n = 1, 2, \dots$$

We now see that the sequence $(|x_n| - |a|)$ is ‘squeezed’ between the null sequence $(x_n - a)$ and the constant null sequence (0) .

Thus, by the Squeeze Rule for sequences, statement (*) holds.

Hence f is continuous.

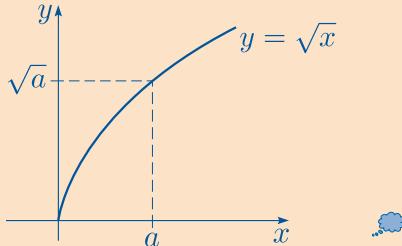
Worked Exercise D46

Determine whether the following function is continuous.

$$f(x) = \sqrt{x} \quad (x \in [0, \infty))$$

Solution

It certainly seems from the graph that f is continuous on its domain, $[0, \infty)$.



We guess that f is continuous on its domain $[0, \infty)$.

Let $a \in [0, \infty)$, and let (x_n) be any sequence in $[0, \infty)$ with $x_n \rightarrow a$ as $n \rightarrow \infty$. We want to prove that

$$x_n \rightarrow a \implies \sqrt{x_n} \rightarrow \sqrt{a},$$

or, in other words, that

$$(x_n - a) \text{ is a null sequence} \implies (\sqrt{x_n} - \sqrt{a}) \text{ is a null sequence. (*)}$$

Now, since $(x_n - a)$ is null, the sequence $(|x_n - a|)$ is null and hence, by the Power Rule for sequences, the sequence $(\sqrt{|x_n - a|})$ is also null.

Next we use an inequality proved in Subsection 3.2 of Unit D1, namely that if $a \geq 0$ and $b \geq 0$, then $\sqrt{|a - b|} \geq |\sqrt{a} - \sqrt{b}|$.

Moreover, since $x_n \geq 0$ and $a \geq 0$, it follows that

$$\sqrt{|x_n - a|} \geq |\sqrt{x_n} - \sqrt{a}|, \quad \text{for } n = 1, 2, \dots$$

We now see that the sequence $(\sqrt{x_n} - \sqrt{a})$ is ‘squeezed’ between the null sequence $(\sqrt{|x_n - a|})$ and the constant null sequence (0) .

Thus, by the Squeeze Rule for sequences, statement (*) holds.

Hence f is continuous.

2.2 Rules for continuous functions

In the previous subsection you saw how to use the definition to check whether a given function is continuous at a point. You will now meet a number of rules that can be used to show that many functions are continuous without the need to go back to the definition. The first set of rules is called the Combination Rules.

Theorem D40 Combination Rules for continuous functions

If f and g are continuous at a , then so are

Sum Rule $f + g$

Multiple Rule λf , for $\lambda \in \mathbb{R}$

Product Rule fg

Quotient Rule f/g , provided that $g(a) \neq 0$.

Proof The proofs of these rules are similar and depend on the corresponding results for sequences. We prove only the Sum Rule.

Suppose that f and g are continuous at a . We want to deduce that $f + g$ is continuous at a .

Let the domain of f be A and the domain of g be B . Then the domain of $f + g$ is $A \cap B$ and this set contains a .

Thus, since

$$(f + g)(x) = f(x) + g(x), \quad \text{for } x \in A \cap B,$$

we have to show that

for each sequence (x_n) in $A \cap B$ such that $x_n \rightarrow a$, we have

$$f(x_n) + g(x_n) \rightarrow f(a) + g(a). \quad (1)$$

We know that (x_n) lies in A and in B , and that both functions f and g are continuous at a . Hence

$$f(x_n) \rightarrow f(a) \quad \text{and} \quad g(x_n) \rightarrow g(a),$$

so statement (1) follows by the Sum Rule for sequences. ■

By using the Combination Rules we can show, for example, that the function $f(x) = 1 - 2x$ is continuous (on \mathbb{R}) since

$$x \mapsto x \text{ is continuous (by Exercise D60(b))}$$

and so

$$x \mapsto -2x \text{ is continuous by the Product Rule.}$$

Also

$$x \mapsto 1 \text{ is continuous (by Exercise D60(a))}$$

and so

$x \mapsto 1 - 2x$ is continuous by the Sum Rule.

Indeed any polynomial $p(x) = a_0 + a_1x + \cdots + a_nx^n, x \in \mathbb{R}$, is continuous at all points of \mathbb{R} since we can build up the expression for p by successive applications of the Combination Rules.

Moreover, it then follows from the Quotient Rule that any rational function $r(x) = p(x)/q(x)$, where p and q are polynomials, is continuous at all points of its domain, that is, all of \mathbb{R} except for points where $q(x) = 0$. We have thus established the following result.

Theorem D41

The following functions are continuous:

- any polynomial $p(x) = a_0 + a_1x + \cdots + a_nx^n$ (on its domain \mathbb{R})
- any rational function $r(x) = p(x)/q(x)$, where p and q are polynomials (on its domain $\mathbb{R} - \{x : q(x) = 0\}$).

The Combination Rules are natural analogues of the corresponding results for sequences. However, we can combine functions in more ways than we can combine sequences: for example, we can *compose* functions f and g to obtain the function $g \circ f$. Composing functions also enables us to obtain ‘new continuous functions from old’, as the next rule shows.

Theorem D42 Composition Rule for continuous functions

If f is continuous at a and g is continuous at $f(a)$, then $g \circ f$ is continuous at a .

Proof Suppose that f is continuous at a and g is continuous at $f(a)$. We want to deduce that $g \circ f$ is continuous at a .

If f has domain A and g has domain B , then the domain of $g \circ f$ is

$$C = \{x \in A : f(x) \in B\}$$

and this set contains a .

Thus we have to show that

for each sequence (x_n) in C such that $x_n \rightarrow a$, we have

$$g(f(x_n)) \rightarrow g(f(a)). \quad (2)$$

We know that (x_n) lies in A and that f is continuous at a , so $f(x_n) \rightarrow f(a)$. Moreover, because (x_n) lies in C , we also know that $(f(x_n))$ lies in B , and since g is continuous at $f(a)$, it follows that $g(f(x_n)) \rightarrow g(f(a))$. Hence statement (2) is true. ■

Worked Exercise D47

Determine whether the following function is continuous.

$$h(x) = \sqrt{x^2 + 1} \quad (x \in \mathbb{R})$$

Solution

We guess that h is continuous on \mathbb{R} .

If we let $f(x) = x^2 + 1$ and $g(x) = \sqrt{x}$, so that $g(f(x)) = \sqrt{x^2 + 1}$, we see that we can express h as a composite function $h = g \circ f$.

Now, f is continuous (on \mathbb{R}), since it is a polynomial, and all its values are positive. Also, g is continuous on $[0, \infty)$, by Worked Exercise D46. It then follows from the Composition Rule that $h = g \circ f$ is continuous.

Exercise D61

Determine whether the following function is continuous.

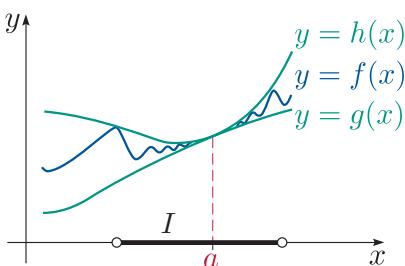
$$f(x) = |x^5| \quad (x \in [0, \infty))$$

Exercise D62

Prove that the following function is continuous, stating each rule or fact that you use.

$$f(x) = \sqrt{x^2 + 2x + 2} - \frac{3x}{x^4 + 4} \quad (x \in \mathbb{R})$$

Next, just as we had a Squeeze Rule for convergent sequences, we show that there is a corresponding Squeeze Rule for continuous functions. The graphs of the functions in the next theorem are illustrated in Figure 18.

**Theorem D43 Squeeze Rule for continuous functions**

Let f , g and h be defined on an open interval I and let $a \in I$. If

1. $g(x) \leq f(x) \leq h(x)$, for $x \in I$,
 2. $g(a) = f(a) = h(a)$, and
 3. g and h are continuous at a ,
- then f is also continuous at a .

Figure 18 The functions in the Squeeze Rule

Proof Suppose that f , g and h satisfy the conditions of the theorem.
We want to prove that f is continuous at a .

Thus we have to show that

for each sequence (x_n) in the domain of f such that $x_n \rightarrow a$, we have
 $f(x_n) \rightarrow f(a)$. (3)

Since $x_n \rightarrow a$ and I is an open interval, there is an integer N such that

$$x_n \in I, \quad \text{for } n > N.$$

Hence, by condition 1,

$$g(x_n) \leq f(x_n) \leq h(x_n), \quad \text{for } n > N.$$

By conditions 2 and 3,

$$\lim_{n \rightarrow \infty} g(x_n) = \lim_{n \rightarrow \infty} h(x_n) = f(a),$$

so statement (3) follows, by the Squeeze Rule for sequences. ■

In the following worked exercise, the only property of the sine function that we need is that $|\sin x| \leq 1$, for $x \in \mathbb{R}$. (We will investigate whether the trigonometric functions sin, cos and tan are continuous in the next subsection.)

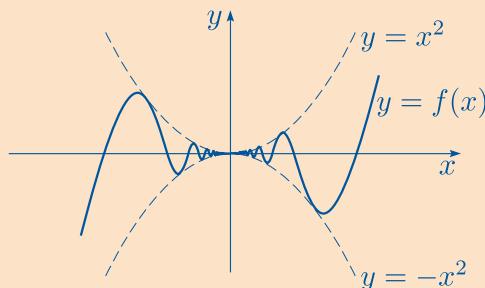
Worked Exercise D48

Determine whether the following function is continuous at 0.

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Solution

The graph of f is shown below.



This suggests that we can find functions g and h that squeeze f near 0, and then use the Squeeze Rule. 💡

We use the Squeeze Rule.

We know that

$$-1 \leq \sin(1/x) \leq 1, \quad \text{for } x \neq 0.$$

Since $x^2 \geq 0$, it follows that

$$-x^2 \leq x^2 \sin(1/x) \leq x^2, \quad \text{for } x \neq 0.$$

Thus, since $f(0) = 0$, we have

$$-x^2 \leq f(x) \leq x^2, \quad \text{for } x \in \mathbb{R}.$$

If we now take $I = \mathbb{R}$, with

$$g(x) = -x^2 \quad \text{and} \quad h(x) = x^2,$$

then

$$g(x) \leq f(x) \leq h(x), \quad \text{for } x \in I,$$

so condition 1 of the Squeeze Rule is satisfied.

Next, $f(0) = g(0) = h(0) = 0$, so condition 2 of the Squeeze Rule is satisfied.

Finally, the functions g and h are polynomials, and so in particular they are continuous at 0. Thus condition 3 of the Squeeze Rule is satisfied.

It then follows from the Squeeze Rule that f is continuous at 0, as required.

Exercise D63

Determine whether the following functions are continuous at 0.

$$(a) \quad f(x) = \begin{cases} x^2 \cos(1/x^2), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

$$(b) \quad f(x) = \begin{cases} \sin(1/x), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Hint: Use the fact that $\sin((2n + \frac{1}{2})\pi) = 1$, for $n \in \mathbb{Z}$.

We now describe another rule for proving that a function is continuous at a point. Consider the hybrid function

$$f(x) = \begin{cases} 1 - x, & x < 0, \\ 0, & 0 \leq x \leq 1, \\ 3x - 3, & x > 1. \end{cases} \quad (4)$$

The domain of f is the whole of \mathbb{R} and the graph of f is shown in Figure 19.

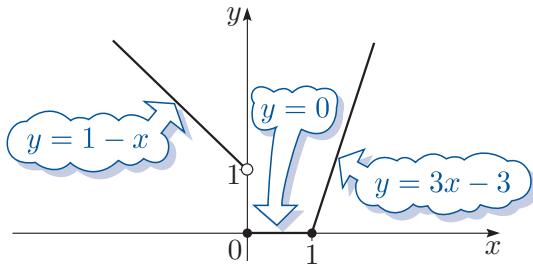


Figure 19 The graph of the hybrid function $y = f(x)$

From the graph, it appears that

1. f is discontinuous at $a = 0$
2. f is continuous at all other values of a .

We can prove that f is discontinuous at 0 by using Strategy D14. We need to find *one* sequence (x_n) such that

$$x_n \rightarrow 0 \quad \text{but} \quad f(x_n) \not\rightarrow f(0).$$

Since $f(0)$ is defined using the rule for $[0, 1]$, we choose a simple sequence (x_n) which tends to 0 from the left; we can choose

$$x_n = -\frac{1}{n}, \quad n = 1, 2, \dots$$

This is illustrated in Figure 20.

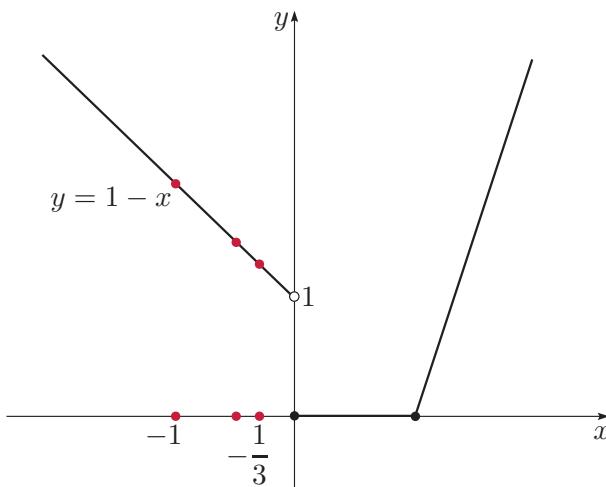


Figure 20 A sequence whose images under f do not tend to $f(0)$

The rule for $f(x)$ which applies for $x < 0$ is $1 - x$, so

$$f(x_n) = f\left(-\frac{1}{n}\right) = 1 - \left(-\frac{1}{n}\right) = 1 + \frac{1}{n}, \quad \text{for } n = 1, 2, \dots,$$

and we also have $f(0) = 0$. Hence

$$x_n \rightarrow 0 \quad \text{but} \quad f(x_n) \rightarrow 1 \neq f(0).$$

Thus f is discontinuous at 0.

But how can we prove that f is continuous at $a = 1$, as the graph suggests? If we define

$$g(x) = 0 \quad (x \in \mathbb{R}) \quad \text{and} \quad h(x) = 3x - 3 \quad (x \in \mathbb{R}),$$

then near the point 1, the graph of f consists of part of the graph $y = g(x)$ to the left of 1, glued to part of the graph $y = h(x)$ to the right of 1.

This idea is the basis of the Glue Rule which is illustrated in Figure 21.

Theorem D44 Glue Rule for continuous functions

Let f be defined on an open interval I and let $a \in I$.

If there are functions g and h such that

1. $f(x) = g(x)$, for $x \in I$, $x < a$,
 $f(x) = h(x)$, for $x \in I$, $x > a$,
 2. $f(a) = g(a) = h(a)$, and
 3. g and h are continuous at a ,
- then f is also continuous at a .

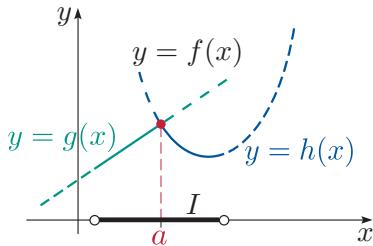


Figure 21 The Glue Rule

Proof Suppose that f , g and h satisfy the conditions of the theorem. We want to prove that

for each sequence (x_n) in the domain of f such that $x_n \rightarrow a$, we have

$$f(x_n) \rightarrow f(a). \quad (5)$$

Since $x_n \rightarrow a$ and I is an open interval, there is an integer N such that

$$x_n \in I, \quad \text{for } n \geq N.$$

Then $(x_n)^\infty_N$ (that is, the sequence x_N, x_{N+1}, \dots) consists of two subsequences (x_{m_k}) and (x_{n_k}) defined by the conditions

$$x_{m_k} < a, \quad \text{for } k = 1, 2, \dots, \quad \text{and} \quad x_{n_k} \geq a, \quad \text{for } k = 1, 2, \dots$$

By conditions 1 and 3, we have

$$g(x_{m_k}) \rightarrow g(a) \quad \text{as } k \rightarrow \infty \quad \text{and} \quad h(x_{n_k}) \rightarrow h(a) \quad \text{as } k \rightarrow \infty.$$

Hence, by conditions 1 and 2, we have

$$f(x_{m_k}) \rightarrow f(a) \quad \text{as } k \rightarrow \infty \quad \text{and} \quad f(x_{n_k}) \rightarrow f(a) \quad \text{as } k \rightarrow \infty.$$

We have shown that the sequence $(f(x_n))$ consists of two subsequences with the same limit, $f(a)$. Recall that, by Theorem D21 in Unit D2, it follows that the whole sequence converges to this limit.

Since both subsequences of $(f(x_n))$ tend to $f(a)$, statement (5) follows. ■

Note that the Glue Rule does not require the functions g and h to be defined on the whole of I , though this is often the case.

We now apply the Glue Rule to the function that we were looking at earlier.

Worked Exercise D49

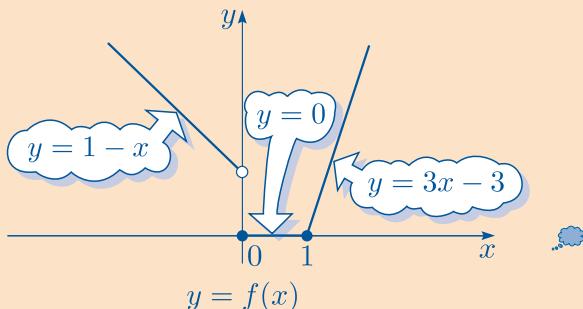
Use the Glue Rule to prove that the function

$$f(x) = \begin{cases} 1 - x, & x < 0, \\ 0, & 0 \leq x \leq 1, \\ 3x - 3, & x > 1, \end{cases}$$

is continuous at 1.

Solution

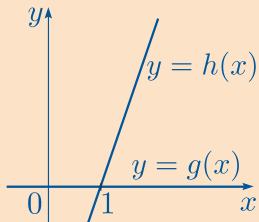
The graph of f is shown again below.



Let I be the open interval $(0, \infty)$ and define the functions

$$g(x) = 0 \quad (x \in \mathbb{R}) \quad \text{and} \quad h(x) = 3x - 3 \quad (x \in \mathbb{R}).$$

The graphs of g and h are shown below.



We chose $I = (0, \infty)$ because on this interval the rule of f is given by the rule of g to the left of 1 and by the rule of h to the right of 1. We could have chosen any smaller open interval that contains the point 1.

Then f is defined on I and $1 \in I$. Also,

$$f(x) = g(x), \quad \text{for } x \in (0, 1),$$

$$f(x) = h(x), \quad \text{for } x \in (1, \infty),$$

so condition 1 of the Glue Rule holds with $a = 1$.

Moreover, $f(1) = g(1) = h(1) = 0$, so condition 2 holds.

Finally, g and h are both polynomials and are therefore both continuous at 1, so condition 3 holds.

Hence f is continuous at 1, by the Glue Rule.

Exercise D64

Prove that the following function is continuous at 1.

$$f(x) = \begin{cases} x^3 - 3x + 5, & x < 1, \\ \frac{2x+1}{3x-2}, & x \geq 1. \end{cases}$$

There are two further straightforward situations in which we can obtain ‘new continuous functions from old’ and we illustrate these by examples. When such situations arise, we normally use the results illustrated here without explicitly referring to them.

1. Consider the function

$$f(x) = \begin{cases} 1-x, & x < 0, \\ 0, & 0 \leq x \leq 1, \\ 3x-3, & x > 1, \end{cases}$$

which we studied in Worked Exercise D49. There we used the Glue Rule to show that f is continuous at 1. We also showed earlier that f is discontinuous at 0. It seems evident that this function f is continuous at all other points in \mathbb{R} . For example, the continuity of f at -1 depends only on the values taken by the function f near the point -1 , and these values are the same as those of the function

$$g(x) = 1-x \quad (x \in \mathbb{R}),$$

which is continuous since it is a polynomial. Since g is continuous at -1 , we deduce that f is also continuous at -1 . A similar argument can be used to show that f is continuous at all points $a \in \mathbb{R} - \{0, 1\}$. Since we have already shown by the Glue Rule that f is continuous at 1, we conclude that f is continuous on $\mathbb{R} - \{0\}$, as expected.

Notice that the above argument works because continuity at a point is a *local property*; that is, it depends only on the behaviour of the function near that point.

2. Consider the function

$$f(x) = x^2 \quad (x \in [-1, 1]).$$

The domain of this function is $[-1, 1]$, and it certainly appears from Figure 22 that f is continuous at each point of $[-1, 1]$. After all, the function

$$g(x) = x^2 \quad (x \in \mathbb{R})$$

is a basic continuous function (since it is a polynomial).

Recall from Unit A1 that if the domain A of a function h is a subset of the domain of a function k and $h(x) = k(x)$ for all $x \in A$, then h is

called the **restriction** of k to A . Thus here f is the restriction of g to the set $[-1, 1]$.

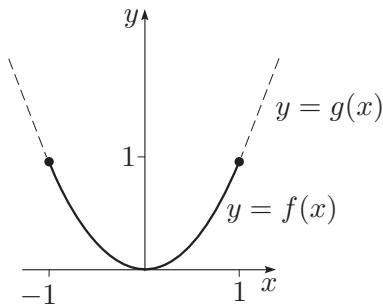


Figure 22 The graph of the restriction of $y = x^2$ to $[-1, 1]$

If we again note that continuity is a local property, it is easy to see from the definition of continuity that if a function f is the restriction of another function g , and g is continuous, then f is also continuous. We can sum up what we have shown as follows:

The restriction of a continuous function is continuous.

2.3 Trigonometric functions and the exponential function

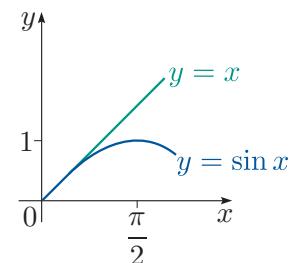
We now prove that the trigonometric functions \sin , \cos and \tan and the exponential function are continuous at all points of \mathbb{R} where they are defined.

Trigonometric functions

We start with a basic inequality for the sine function. This is illustrated in Figure 23.

Theorem D45 Sine Inequality

$$\sin x \leq x, \quad \text{for } 0 \leq x \leq \frac{\pi}{2}.$$



Proof If $x = 0$, then $\sin 0 = 0$, so there is equality.

Suppose next that $0 < x \leq \pi/2$, and consider the diagram in Figure 24, which represents a quarter circle, centred at the origin, with radius 1.

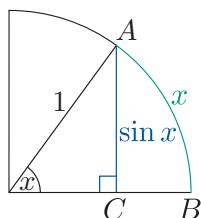


Figure 24 An arc of a circle of length x

Figure 23 The graphs of $y = \sin x$ and $y = x$

Since the circle has radius 1, the arc AB has length x equal to the angle x (measured in radians) and the perpendicular AC has length $\sin x$. Hence

$$\sin x < x, \quad \text{for } 0 < x \leq \frac{\pi}{2}.$$

Combining this inequality with the fact that $\sin 0 = 0$ gives the result. ■

We can now deduce the crucial inequality for proving the continuity of the sine function. This is illustrated in Figure 25.

Corollary D46

$$|\sin x| \leq |x|, \quad \text{for } x \in \mathbb{R}.$$

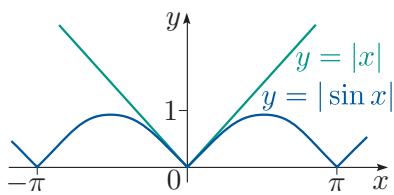


Figure 25 The graphs of $y = |\sin x|$ and $y = |x|$

Proof The Sine Inequality shows that this inequality holds for $0 \leq x \leq \pi/2$. For $x > \pi/2$, we have

$$|\sin x| \leq 1 < \frac{\pi}{2} < x = |x|,$$

so the inequality is also true in this case.

Finally, the inequality holds for $x < 0$, since

$$|\sin(-x)| = |\sin x| \quad \text{and} \quad |-x| = |x|.$$

This is the key tool that we need to prove the continuity of the trigonometric functions.

Theorem D47

The trigonometric functions sine, cosine and tangent are continuous.

Proof To prove that the sine function is continuous at $a \in \mathbb{R}$, we want to show that

for each sequence (x_n) in \mathbb{R} such that $x_n \rightarrow a$, we have

$$\sin x_n \rightarrow \sin a. \tag{6}$$

To do this, we use the trigonometric identity

$$\sin x - \sin a = 2 \cos\left(\frac{1}{2}(x + a)\right) \sin\left(\frac{1}{2}(x - a)\right).$$

This identity can be obtained by writing

$$x = \frac{1}{2}(x + a) + \frac{1}{2}(x - a)$$

and

$$a = \frac{1}{2}(x + a) - \frac{1}{2}(x - a),$$

and then using the formulas for $\sin(A + B)$ and $\sin(A - B)$ which you can find in the module Handbook. ☀

We obtain

$$\begin{aligned} |\sin x_n - \sin a| &= |2 \cos\left(\frac{1}{2}(x_n + a)\right) \sin\left(\frac{1}{2}(x_n - a)\right)| \\ &\leq 2 |\sin\left(\frac{1}{2}(x_n - a)\right)| \quad (\text{since } |\cos x| \leq 1) \\ &\leq 2 \left|\frac{1}{2}(x_n - a)\right| \quad (\text{by Corollary D46}) \\ &= |x_n - a|. \end{aligned}$$

Thus, if $(x_n - a)$ is null, then $(\sin x_n - \sin a)$ is null, by the Squeeze Rule for sequences, so statement (6) holds.

The continuity of the cosine and tangent functions now follows from the identities

$$\cos x = \sin\left(x + \frac{1}{2}\pi\right) \quad \text{and} \quad \tan x = \frac{\sin x}{\cos x},$$

using the Composition Rule and the Quotient Rule. ■

Exercise D65

Prove that the following function is continuous (on \mathbb{R}), stating each rule or fact about continuity that you use.

$$f(x) = x^2 + 1 + 3 \sin\left(\sqrt{x^2 + 1}\right).$$

The exponential function

We start with two fundamental inequalities for the exponential function. These are illustrated in Figure 26.

Theorem D48 Exponential Inequalities

- (a) $e^x \geq 1 + x$, for $x \geq 0$
- (b) $e^x \leq \frac{1}{1-x}$, for $0 \leq x < 1$.

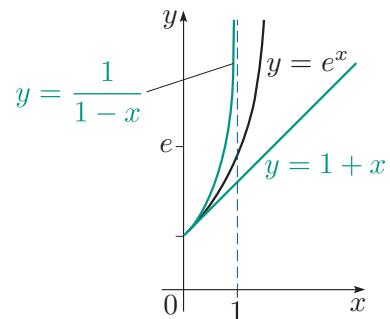


Figure 26 The graphs of $y = e^x$, $y = 1/(1-x)$ and $y = 1 + x$

Proof We prove both inequalities using the exponential series which you met in Unit D3 *Series*:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots, \quad \text{for } x \geq 0.$$

- (a) For $x \geq 0$, we have $x^2/2! \geq 0$, $x^3/3! \geq 0$, and so on. Hence

$$e^x \geq 1 + x, \quad \text{for } x \geq 0.$$

- (b) For $x \geq 0$, we also have $x^2/2! \leq x^2$, $x^3/3! \leq x^3$, and so on. Hence

$$e^x \leq 1 + x + x^2 + x^3 + \dots.$$

The series on the right is a geometric series, which is convergent with sum $1/(1-x)$, for $0 \leq x < 1$. Hence

$$e^x \leq \frac{1}{1-x}, \quad \text{for } 0 \leq x < 1. \quad \blacksquare$$

We can now deduce the following inequalities which we use in proving the continuity of the exponential function. These are illustrated in Figure 27.

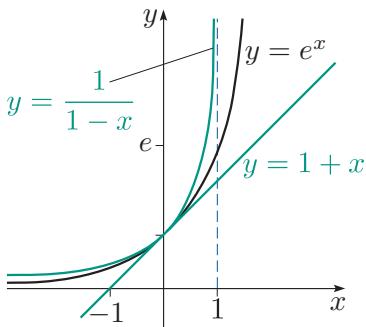


Figure 27 The graphs of the functions in Corollary D49

Corollary D49

$$1+x \leq e^x \leq \frac{1}{1-x}, \quad \text{for } |x| < 1.$$

Proof The Exponential Inequalities show that these inequalities both hold for $0 \leq x < 1$. For $-1 < x < 0$, we have $0 < -x < 1$, so

$$1+(-x) \leq e^{-x} \leq \frac{1}{1-(-x)}.$$

Taking reciprocals and reversing these inequalities (which is possible since all three expressions are positive for $-1 < x < 0$), we obtain

$$1+x \leq e^x \leq \frac{1}{1-x}, \quad \text{for } -1 < x < 0.$$

Hence

$$1+x \leq e^x \leq \frac{1}{1-x}, \quad \text{for } |x| < 1. \quad \blacksquare$$

We can now prove the continuity of the exponential function.

Theorem D50

The exponential function is continuous.

Proof To prove that the exponential function is continuous at $a \in \mathbb{R}$, we want to show that

for each sequence (x_n) in \mathbb{R} such that $x_n \rightarrow a$, we have

$$e^{x_n} \rightarrow e^a. \quad (7)$$

Now, if $(x_n - a)$ is null, then there is a positive integer N such that $|x_n - a| < 1$, for $n > N$. Applying Corollary D49, with $x_n - a$ instead of x , we obtain

$$1 + (x_n - a) \leq e^{x_n - a} \leq \frac{1}{1 - (x_n - a)}, \quad \text{for } n > N.$$

Thus $e^{x_n - a} \rightarrow 1$, by the Squeeze Rule for sequences. Hence $e^{x_n} = e^a e^{x_n - a} \rightarrow e^a$, so statement (7) holds. \blacksquare

Exercise D66

Prove that the following function is continuous, stating each rule or fact about continuity that you use.

$$f(x) = x^5 - 5x^2 + 7e^{-x^2}$$

We end this subsection with a reminder of the various approaches that you have met for investigating the continuity of a function $f : A \rightarrow \mathbb{R}$ at $a \in A$. Recall that you should first guess whether f is continuous or discontinuous at a , then check whether your guess is correct (a sketch of the graph may help you make your guess). You can check your guess using Strategy D14. You have also seen that, in many cases, it is possible to show that f is continuous at a by applying rules such as the Combination Rules, the Composition Rule, the Squeeze Rule and the Glue Rule to functions which you already know to be continuous. We have proved that a number of familiar functions are continuous and we now collect these together in the following result.

Theorem D51 Basic continuous functions

The following functions are continuous:

- polynomials and rational functions
- $f(x) = |x|$
- $f(x) = \sqrt{x}$
- the trigonometric functions sine, cosine and tangent
- the exponential function.

3 Properties of continuous functions

In this section you will meet some of the fundamental properties of continuous functions, and see that these properties hold for continuous functions defined on **bounded closed intervals**; that is, intervals of the form $[a, b]$. You will also see some applications of these properties, in particular to locating the zeros of a continuous function.

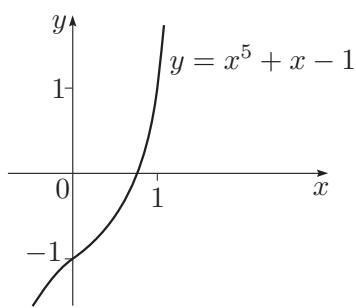


Figure 28 The graph of $y = x^5 + x - 1$

3.1 Intermediate Value Theorem

At the end of Section 1 we considered the function

$$f(x) = x^5 + x - 1 \quad (x \in \mathbb{R})$$

and we pointed out that it is not easy to prove that $f(\mathbb{R}) = \mathbb{R}$.

For example, is there a value of x such that $f(x) = 0$? In other words, is there a solution of the equation

$$x^5 + x - 1 = 0?$$

The shape of the graph of $y = x^5 + x - 1$ shown in Figure 28 certainly suggests that such a number x exists. Since $f(0) = -1$ and $f(1) = 1$, we expect there to be some number x in the interval $(0, 1)$ such that $f(x) = 0$. However, we do not have a formula for solving the above equation to find x .

The key to showing that such a number x exists lies in the fact that f is a continuous function, so there cannot be any gaps in its graph; this is the essence of the *Intermediate Value Theorem*. We prove this result at the end of this subsection but first we show how it can be used.

Theorem D52 Intermediate Value Theorem

Let f be a function continuous on $[a, b]$ such that $f(a) \neq f(b)$, and let k be any number lying between $f(a)$ and $f(b)$. Then there exists a number c in (a, b) such that

$$f(c) = k.$$

The first purely analytic proof of the Intermediate Value Theorem was given by the Bohemian mathematician and theologian Bernard Bolzano (1781–1848) in a paper published in 1817. Bolzano was one of the first mathematicians to begin to instil rigour into analysis. To quote Steve Russ, the leading expert on Bolzano's mathematics, this paper

‘represents an important stage in the rigorous foundation of analysis and is one of the earliest occasions when the continuity of a function and the convergence of an infinite series are both defined and used correctly.’

Four years later, the theorem appeared in the *Cours d'Analyse* of Augustin-Louis Cauchy (1789–1857), but it is unlikely that Cauchy knew of Bolzano's work.

(Source: Russ, S.B. (1980) ‘A Translation of Bolzano's Paper on the Intermediate Value Theorem’, *Historia Mathematica*, vol. 7, no. 2, pp. 156–185.)



Bernard Bolzano

The Intermediate Value Theorem has two possible cases: we have either

$$f(a) < k < f(b) \quad \text{or} \quad f(a) > k > f(b).$$

These are illustrated in Figure 29.

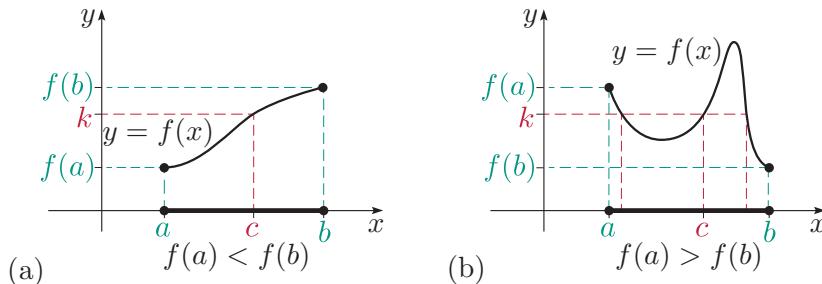


Figure 29 The two cases of the Intermediate Value Theorem

As the graph in Figure 29(b) shows, there may be more than one possible value of c such that $f(c) = k$.

The conclusion of the Intermediate Value Theorem may be *false* if f is discontinuous at even *one* point of $[a, b]$. For example, the function

$$f(x) = \begin{cases} 1/x, & 0 < |x| \leq 1, \\ 0, & x = 0, \end{cases} \quad (8)$$

is continuous on $[-1, 1]$ except at 0. For this function,

$$f(-1) = -1 \quad \text{and} \quad f(1) = 1,$$

but there is no number c in $(-1, 1)$ such that $f(c) = \frac{1}{2}$, as shown in Figure 30. The conclusion of the Intermediate Value Theorem may be also be false if $f(a) = f(b)$. For example, if $f(x) = x^2$ with $a = -1$ and $b = 1$, then $f(a) = f(b) = 1$ and the only possible value of k is 1, but there is no $c \in (-1, 1)$ such that $f(c) = 1$.

Here is a typical application of the Intermediate Value Theorem.

Worked Exercise D50

Use the Intermediate Value Theorem to prove that there is a number c in $(0, 1)$ such that

$$c^5 + c - 1 = 0.$$

Solution

Consider the basic continuous function

$$f(x) = x^5 + x - 1.$$

Then f is continuous on $[0, 1]$ and also

$$f(0) = -1 \quad \text{and} \quad f(1) = 1.$$

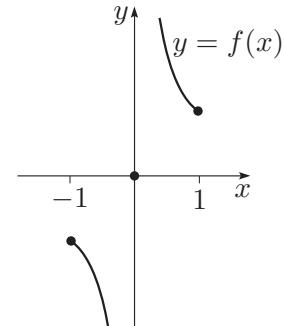


Figure 30 A function to which the Intermediate Value Theorem does not apply

Since $f(0) < 0 < f(1)$, it follows from the Intermediate Value Theorem that there is a number c in $(0, 1)$ such that

$$f(c) = 0; \quad \text{that is, } c^5 + c - 1 = 0.$$

Notice that the function f is strictly increasing on $[0, 1]$, so in this case the number c must be *unique*.

To obtain further information about the location of the number c in Worked Exercise D50, we can use, for example, the **bisection method**. This involves repeatedly bisecting the interval containing the solution and determining the values of f at the bisection points, in order to find shorter and shorter intervals in which the number c must lie.

For example, at the first bisection the function

$$f(x) = x^5 + x - 1$$

satisfies

$$f\left(\frac{1}{2}\right) = \left(\frac{1}{2}\right)^5 + \frac{1}{2} - 1 < 0 \quad \text{and} \quad f(1) = 1 > 0,$$

so the number c must lie in $(\frac{1}{2}, 1)$. To find an interval of length $\frac{1}{4}$ containing c , we next consider the value $f\left(\frac{3}{4}\right)$, and so on.

Exercise D67

Use the bisection method to find an interval of length $\frac{1}{16}$ containing the number c such that

$$c^5 + c - 1 = 0.$$

The bisection method can also be used to prove the Intermediate Value Theorem. We prove the special case of this result when $k = 0$ and $f(a) < f(b)$; the proof when $k = 0$ and $f(a) > f(b)$ is similar. The general case can then be deduced from the case when $k = 0$ by considering the function

$$F(x) = f(x) - k.$$

Theorem D53 Intermediate Value Theorem (special case)

Let f be a function continuous on $[a, b]$ and suppose that

$$f(a) < 0 < f(b).$$

Then there exists a number c in (a, b) such that

$$f(c) = 0.$$

Proof We use the bisection method.

First we define $[a_0, b_0] = [a, b]$ and let $p = \frac{1}{2}(a_0 + b_0)$, the midpoint of $[a_0, b_0]$. If $f(p) = 0$, then the proof is complete, since we can take $c = p$. Otherwise, we define

$$[a_1, b_1] = \begin{cases} [a_0, p], & \text{if } f(p) > 0, \\ [p, b_0], & \text{if } f(p) < 0. \end{cases}$$

In either case, we have

1. $[a_1, b_1] \subseteq [a_0, b_0]$
2. $b_1 - a_1 = \frac{1}{2}(b_0 - a_0)$
3. $f(a_1) < 0 < f(b_1)$.

Now we repeat this process, bisecting $[a_1, b_1]$ to obtain $[a_2, b_2]$, and so on. If, at any stage, we encounter a bisection point p such that $f(p) = 0$, then the proof is complete. Otherwise, we obtain a sequence of closed intervals

$$[a_n, b_n], \quad n = 0, 1, 2, \dots,$$

with the properties that, for $n = 0, 1, 2, \dots$,

1. $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$
2. $b_n - a_n = \left(\frac{1}{2}\right)^n (b_0 - a_0)$
3. $f(a_n) < 0 < f(b_n)$.

Property 1 implies that (a_n) is increasing and bounded above by b_0 .

We now use the Monotone Convergence Theorem (Theorem D22 in Unit D2), which says that any sequence that is increasing and bounded above must be convergent.

Hence, by the Monotone Convergence Theorem, (a_n) is convergent. Let

$$\lim_{n \rightarrow \infty} a_n = c.$$

This is illustrated in Figure 31.

By property 2 and the Combination Rules for sequences,

$$\begin{aligned} \lim_{n \rightarrow \infty} b_n &= \lim_{n \rightarrow \infty} (a_n + (b_n - a_n)) \\ &= \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n (b_0 - a_0) \\ &= c + 0 = c. \end{aligned}$$

Now we use the fact that f is continuous at c to obtain

$$\lim_{n \rightarrow \infty} f(a_n) = f(c) \quad \text{and} \quad \lim_{n \rightarrow \infty} f(b_n) = f(c).$$

By property 3, $f(a_n) < 0$, for $n = 0, 1, 2, \dots$, so $f(c) \leq 0$, by the Limit Inequality Rule (Theorem D11 in Unit D2). Likewise, $f(c) \geq 0$ because $f(b_n) > 0$, for $n = 0, 1, 2, \dots$. Hence $f(c) = 0$, as required. ■

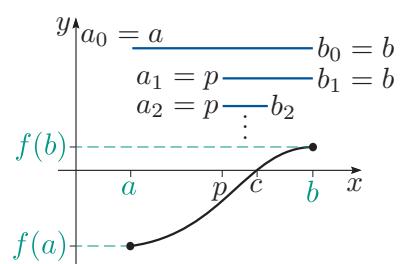


Figure 31 The limit c

3.2 Locating zeros of continuous functions

If f is a function and c is a real number such that

$$f(c) = 0,$$

then c is called a **zero** of the function f . We sometimes say that the function **vanishes** at c .

We often show that an equation has a solution by proving that a related continuous function has a zero (using the Intermediate Value Theorem with $k = 0$). You saw an example of this in Worked Exercise D50 and we give another one now.

Worked Exercise D51

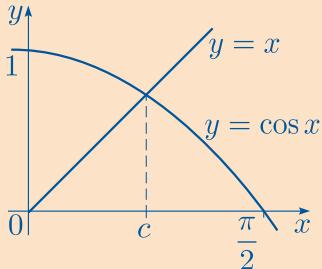
Prove that the equation

$$\cos x = x$$

has a solution in the interval $(0, 1)$.

Solution

The graphs shown below suggest that there is a solution to this equation.



We now prove that this is the case.

We consider the function

$$f(x) = \cos x - x$$

and show that f has a zero c in $(0, 1)$. Now f is continuous, by the Combination Rules. Moreover,

$$f(0) = \cos 0 - 0 = 1 > 0$$

and

$$f(1) = \cos 1 - 1 < 0.$$

Thus, by the Intermediate Value Theorem, there is a number c in $(0, 1)$ such that

$$f(c) = 0, \quad \text{so} \quad \cos c = c.$$

The result that was proved in Worked Exercise D51 is a special case of the result that you are asked to prove in the next exercise.

Exercise D68

Suppose that the function $f : [0, 1] \rightarrow [0, 1]$ is continuous. Prove that the equation

$$f(x) = x$$

has a solution c in the interval $[0, 1]$.

Hint: Consider the function $g(x) = f(x) - x$ ($x \in [0, 1]$).

Zeros of polynomials

We now consider the problem of locating the zeros of polynomial functions. Recall that zeros of polynomials were discussed in Unit A2 *Number systems*, where you met the fact that a polynomial of degree n has at most n zeros.

First try the following exercise.

Exercise D69

Let

$$p(x) = x^6 - 4x^4 + x + 1 \quad (x \in \mathbb{R}).$$

Prove that p has a zero in each of the intervals $(-1, 0)$, $(0, 1)$ and $(1, 2)$.

When we wish to locate the zeros (if any) of a given polynomial, we can begin by applying the following result, which gives an interval in which the zeros *must* lie.

Theorem D54

Let

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \quad (x \in \mathbb{R}),$$

where $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$. Then all the zeros of p (if there are any) lie in the open interval $(-M, M)$, where

$$M = 1 + \max\{|a_{n-1}|, \dots, |a_1|, |a_0|\}.$$

Proof Suppose we write

$$r(x) = \frac{p(x)}{x^n} - 1 = \frac{a_{n-1}}{x} + \cdots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \quad (x \in \mathbb{R} - \{0\}),$$

so that

$$p(x) = x^n(1 + r(x)), \quad \text{for } x \neq 0. \quad (9)$$

Using the Triangle Inequality, we obtain

$$\begin{aligned} |r(x)| &= \left| \frac{a_{n-1}}{x} + \cdots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \right| \\ &\leq \left| \frac{a_{n-1}}{x} \right| + \cdots + \left| \frac{a_1}{x^{n-1}} \right| + \left| \frac{a_0}{x^n} \right| \\ &\leq (M-1) \left(\frac{1}{|x|} + \cdots + \frac{1}{|x|^{n-1}} + \frac{1}{|x|^n} \right). \end{aligned}$$

Thus if $|x| > 1$, so that $1/|x| < 1$, we have

$$\begin{aligned} |r(x)| &< (M-1) \left(\frac{1}{|x|} + \frac{1}{|x|^2} + \cdots \right) \\ &= (M-1) \frac{(1/|x|)}{1 - 1/|x|} = \frac{M-1}{|x|-1}, \end{aligned}$$

by summing the convergent geometric series.

Hence for $|x| \geq M$ we have $|r(x)| < 1$, and therefore $1 + r(x) > 0$. It now follows from equation (9) that $p(x)$ has the same sign as x^n for $|x| \geq M$. Since $x^n \neq 0$ for $|x| \geq M$, any zero of p must lie in the interval $(-M, M)$, as required. ■

The next worked exercise shows how we can use Theorem D54 in combination with the Intermediate Value Theorem to obtain information about the zeros of a polynomial.

Worked Exercise D52

Prove that the following polynomial has at least two zeros.

$$p(x) = 2x^4 - 4x^2 - 2x + 2 \quad (x \in \mathbb{R})$$

Solution

We first note that Theorem D54 can only be applied to polynomials for which the coefficient of x^n is equal to 1. So we find a polynomial that satisfies this condition and has the same zeros as p . 

We have

$$p(x) = 2x^4 - 4x^2 - 2x + 2 = 2(x^4 - 2x^2 - x + 1).$$

For $q(x) = x^4 - 2x^2 - x + 1$ we have

$$M = 1 + \max\{|-2|, |-1|, |1|\} = 3,$$

and so all the zeros of q , and hence of p , lie in $(-3, 3)$, by Theorem D54.

We now compile a table of values of $p(n)$, for integers n in $[-3, 3]$.

n	-3	-2	-1	0	1	2	3
$p(n)$	134	22	2	2	-2	14	122

We find that $p(0)$ and $p(1)$ have opposite signs, as do $p(1)$ and $p(2)$. Thus, since p is continuous, it has a zero in each of the open intervals $(0, 1)$ and $(1, 2)$, by the Intermediate Value Theorem.

Thus we have proved that p has *at least* two zeros.

Since p is a polynomial of degree 4, we also know that it has *at most* 4 zeros. It can be shown that p has *exactly* two zeros, though we do not prove this here.

When using Theorem D54, it is often not necessary to calculate the values of $p(n)$ for *all* the integers n in the interval $[-M, M]$. For example, in Worked Exercise D52, in order to locate two zeros it would have been sufficient to calculate the values of $p(n)$ for n in $[-2, 2]$. Often it is a good idea to calculate the values of $p(n)$ for small values of n first and only calculate $p(n)$ for larger values if you have to; in other words, start filling in a table like that in the solution to Worked Exercise D52 in the middle and work as far outwards as necessary.

Exercise D70

Prove that the following polynomial has at least three zeros.

$$p(x) = x^5 + 3x^4 - x - 1 \quad (x \in \mathbb{R})$$

3.3 Extreme Value Theorem

We now describe another important property of continuous functions. First we give the following definitions.

Definitions

Let f be a function with domain A . Then

- f has **maximum value** $f(c)$ in A if $c \in A$ and

$$f(x) \leq f(c), \quad \text{for } x \in A$$

- f has **minimum value** $f(c)$ in A if $c \in A$ and

$$f(c) \leq f(x), \quad \text{for } x \in A$$

- f is **bounded** on A if, for some $M \in \mathbb{R}$,

$$|f(x)| \leq M, \quad \text{for } x \in A.$$

An **extreme value** is either a maximum or a minimum value.

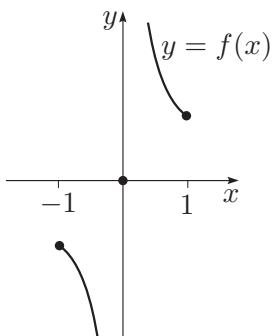


Figure 32 A function that is not bounded on its domain

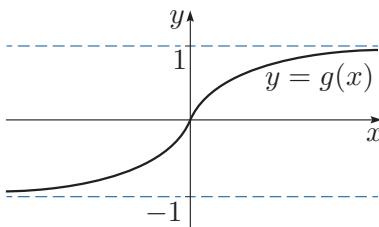


Figure 33 A bounded function with no extreme values

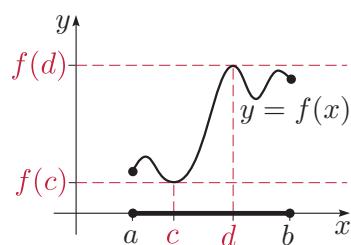


Figure 34 Extreme values of f

For example, the function $f(x) = \sin x$, with domain \mathbb{R} , has maximum value 1 in \mathbb{R} , since

$$\sin x \leq 1 = \sin \frac{\pi}{2}, \quad \text{for } x \in \mathbb{R}.$$

Also, this function f is bounded on \mathbb{R} , since

$$|\sin x| \leq 1, \quad \text{for } x \in \mathbb{R}.$$

On the other hand, the function

$$f(x) = \begin{cases} 1/x, & 0 < |x| \leq 1, \\ 0, & x = 0, \end{cases}$$

is not bounded on its domain $[-1, 1]$; see Figure 32.

Note that although a function that has both a maximum value and a minimum value is bounded, a function can be bounded without having a maximum value or a minimum value; for example, the function

$$g(x) = \frac{x}{1 + |x|}$$

is bounded on its domain \mathbb{R} , since

$$|g(x)| = \frac{|x|}{1 + |x|} < 1, \quad \text{for } x \in \mathbb{R},$$

but g is strictly increasing on \mathbb{R} , so it has no maximum or minimum value on \mathbb{R} . The graph of g is shown in Figure 33.

The following result states that a continuous function on a bounded closed interval always has a maximum value and a minimum value. It is illustrated in Figure 34.

Theorem D55 Extreme Value Theorem

Let f be a function continuous on $[a, b]$. Then there exist numbers c and d in $[a, b]$ such that

$$f(c) \leq f(x) \leq f(d), \quad \text{for } x \in [a, b].$$

An immediate consequence of the Extreme Value Theorem is the following corollary.

Corollary D56 Boundedness Theorem

Let f be a function continuous on $[a, b]$. Then there exists a number M such that

$$|f(x)| \leq M, \quad \text{for } x \in [a, b].$$

The function

$$f(x) = \begin{cases} 1/x, & 0 < |x| \leq 1, \\ 0, & x = 0, \end{cases}$$

whose graph was given in Figure 32, shows that the conclusions of the Extreme Value Theorem and the Boundedness Theorem may be *false* if f is discontinuous at even *one* point of $[a, b]$.

We now prove the Extreme Value Theorem. The proof illustrates the use of many of the techniques that you have met so far but if you are short of time you may prefer to skim through it, noting the main steps.

Proof of the Extreme Value Theorem We prove that there exists a number d in $[a, b]$ such that

$$f(x) \leq f(d), \quad \text{for } x \in [a, b].$$

We use the function

$$g(x) = \frac{x}{1 + |x|},$$

which is strictly increasing and continuous on \mathbb{R} , with

$$|g(x)| < 1, \quad \text{for } x \in \mathbb{R}.$$

The graph $y = g(x)$ was shown in Figure 33.

Then the function

$$h(x) = g(f(x)) \quad (x \in [a, b])$$

is continuous, by the Composition Rule, and

$$|h(x)| < 1, \quad \text{for } x \in [a, b].$$

• The function h is easier to work with than the function f and, since g is strictly increasing, if we can find d such that $h(x) \leq h(d)$ for $x \in (a, b)$, then we can deduce that $f(x) \leq f(d)$ for $x \in [a, b]$.

Hence the image set $h([a, b])$ is bounded. Thus, by the Least Upper Bound Property of \mathbb{R} which you met in Section 4 of Unit D1,

the supremum, M say, of $h([a, b])$ exists.

• Remember that the least upper bound of a set A of real numbers is also called its supremum, and is denoted by $\sup A$.

We now use the bisection method to find $d \in [a, b]$ such that $h(d) = M$.

We define $[a_0, b_0] = [a, b]$ and $p = \frac{1}{2}(a_0 + b_0)$. Then at least one of the image sets $h([a_0, p])$ and $h([p, b_0])$ must have least upper bound M .

• We now choose the interval $[a_1, b_1]$ to be whichever of $h([a_0, p])$ and $h([p, b_0])$ has least upper bound M , or either if both do.

Thus we can choose $[a_1, b_1]$ such that

1. $[a_1, b_1] \subseteq [a_0, b_0]$
2. $b_1 - a_1 = \frac{1}{2}(b_0 - a_0)$
3. $M = \sup h([a_1, b_1]).$

Now we repeat this process, bisecting $[a_1, b_1]$ to obtain $[a_2, b_2]$, and so on. We obtain a sequence of closed intervals

$$[a_n, b_n], \quad n = 0, 1, 2, \dots,$$

with the following properties for $n = 0, 1, 2, \dots$:

1. $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$
2. $b_n - a_n = \left(\frac{1}{2}\right)^n (b_0 - a_0)$
3. $M = \sup h([a_n, b_n]).$

As in the proof of the Intermediate Value Theorem, properties 1 and 2 imply that there is a real number $d \in [a, b]$ such that

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = d.$$

By property 3, for each $n = 1, 2, \dots$, there is a number t_n such that

$$a_n \leq t_n \leq b_n \quad \text{and} \quad M - 1/n \leq h(t_n) \leq M,$$

because each number $M - 1/n$ is *not* an upper bound of the image set $h([a, b])$.

Hence, by the Squeeze Rule for sequences,

$$\lim_{n \rightarrow \infty} t_n = d \quad \text{and} \quad \lim_{n \rightarrow \infty} h(t_n) = M.$$

Thus, by the continuity of h at d ,

$$h(d) = \lim_{n \rightarrow \infty} h(t_n) = M,$$

so

$$h(x) = g(f(x)) \leq g(f(d)) = h(d), \quad \text{for } x \in [a, b].$$

Since g is strictly increasing, it follows that

$$f(x) \leq f(d), \quad \text{for } x \in [a, b],$$

as required.

Similar reasoning shows that there exists $c \in [a, b]$ such that

$$f(c) \leq f(x), \quad \text{for } x \in [a, b].$$



Antipodal Points Theorem (optional)

We conclude this section by presenting a corollary of the Intermediate Value Theorem. This application is included for your interest and is not assessed.

We ask whether there must always be a pair of *antipodal points* (that is, points which lie at opposite ends of a line segment through the centre of the Earth) on the equator at which the temperature is the same. (The equator is shown in Figure 35.) The following result uses the Intermediate Value Theorem to show that the answer is ‘yes’. Here $g(\theta)$ represents the temperature at a point on the equator at an angle θ radians east of the Greenwich meridian. If $g(c) = g(c + \pi)$, then c and $c + \pi$ represent antipodal points with the same temperature.



Figure 35 The equator

Theorem D57 Antipodal Points Theorem

If $g : [0, 2\pi] \rightarrow \mathbb{R}$ is a continuous function and $g(0) = g(2\pi)$, then there exists a number c in $[0, \pi]$ such that

$$g(c) = g(c + \pi).$$

Proof The theorem is illustrated in Figure 36 below.

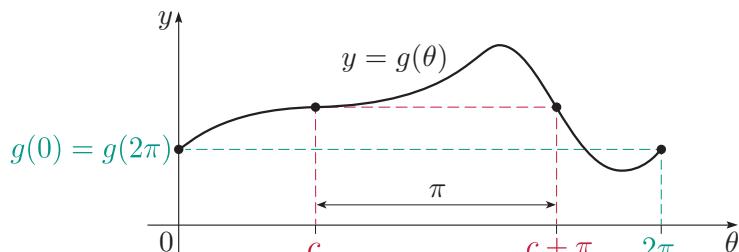


Figure 36 The Antipodal Points Theorem

First note that if $g(0) = g(\pi)$, then we can take $c = 0$. So let us assume that

$$g(0) < g(\pi). \tag{10}$$

(The proof in the case $g(0) > g(\pi)$ is similar.)

Now we define

$$h(\theta) = g(\theta + \pi) \quad (\theta \in [0, \pi]),$$

and consider the graphs $y = g(\theta)$ and $y = h(\theta)$, for $0 \leq \theta \leq \pi$, shown in Figure 37. The graph $y = h(\theta)$ is obtained by translating to the left the part of the graph $y = g(\theta)$ corresponding to $\pi \leq \theta \leq 2\pi$.

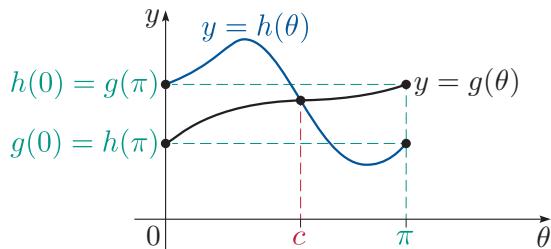


Figure 37 The graphs of g and h

Since

$$h(0) = g(0 + \pi) = g(\pi) \quad \text{and} \quad h(\pi) = g(\pi + \pi) = g(0),$$

inequality (10) can be rewritten as

$$g(0) < h(0) \quad \text{and} \quad g(\pi) > h(\pi).$$

This suggests that our two graphs must cross at some point c in $(0, \pi)$, giving

$$g(c) = h(c) \quad \text{and hence} \quad g(c) = g(c + \pi).$$

To make this argument rigorous, we define a function f as

$$f(\theta) = g(\theta) - h(\theta) \quad (\theta \in [0, \pi]),$$

which is continuous on $[0, \pi]$. Also,

$$f(0) = g(0) - h(0) < 0 \quad \text{and} \quad f(\pi) = g(\pi) - h(\pi) > 0.$$

Thus, by the Intermediate Value Theorem with $k = 0$, there exists a number c in $(0, \pi)$ such that $f(c) = 0$, so $g(c) = h(c)$ and hence

$$g(c) = g(c + \pi),$$

as required. ■



Karol Borsuk



Stanislaw Ulam

The Antipodal Points Theorem is the one-dimensional case of the Borsuk–Ulam theorem, an important result about continuous functions that holds in any number of dimensions. It is named after the Polish mathematicians Karol Borsuk (1905–1982) and Stanisław Ulam (1909–1984) – Ulam was the first to formulate the theorem and in 1933 Borsuk was the first to prove it. The two-dimensional case of the theorem can be illustrated by saying that at any moment there is always pair of antipodal points on the Earth’s surface with equal temperatures and barometric pressures.

4 Inverse functions

In this section you will meet the Inverse Function Rule, which gives conditions for a continuous function f to have a continuous inverse function f^{-1} . You will then see how the Inverse Function Rule can be used to show that the inverse functions of various standard functions exist and are continuous. Finally, you will see how to use the exponential function and its inverse function to define a^x , for $a > 0$ and all $x \in \mathbb{R}$.

4.1 Inverse Function Rule

At the end of Section 1 we discussed the function

$$f(x) = x^5 + x - 1 \quad (x \in \mathbb{R}).$$

We showed that f is strictly increasing and hence one-to-one, but we could not prove that $f(\mathbb{R}) = \mathbb{R}$, so we could not prove that the inverse function f^{-1} has domain \mathbb{R} . In this section you will see that f^{-1} does indeed have domain \mathbb{R} and, moreover, f^{-1} is continuous on \mathbb{R} . We now know that f is continuous (as it is a polynomial) and this means that we can apply a result known as the *Inverse Function Rule*. The proof of this result is based on the Intermediate Value Theorem and is given in Subsection 4.4. The graphs of the functions in the statement of the theorem are shown in Figures 38 and 39.

Theorem D58 Inverse Function Rule

Let $f : I \rightarrow J$, where I is an interval and J is the image set $f(I)$, be a function such that

1. f is strictly increasing on I
2. f is continuous on I .

Then J is an interval and f has an inverse function $f^{-1} : J \rightarrow I$ such that

- 1'. f^{-1} is strictly increasing on J
- 2'. f^{-1} is continuous on J .

Remarks

1. The interval I can be *any* type of interval: open or closed, half-open, bounded or unbounded.
2. There is another version of the Inverse Function Rule with ‘strictly increasing’ replaced by ‘strictly decreasing’.

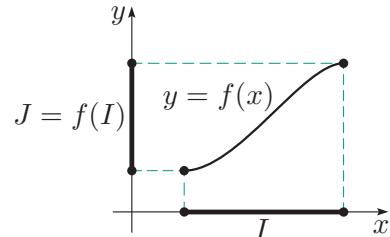


Figure 38 The graph of $f : I \rightarrow J$

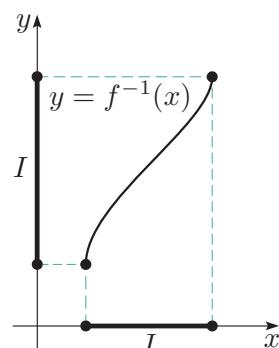


Figure 39 The graph of $f^{-1} : J \rightarrow I$

If we return to the function

$$f(x) = x^5 + x - 1 \quad (x \in \mathbb{R}),$$

then this satisfies the conditions of the Inverse Function Rule with $I = \mathbb{R}$. So we can deduce that f has a *continuous* inverse function and that the domain of this inverse function is an interval $J = f(I) = f(\mathbb{R})$, but we still need to determine what this interval is.

In general, in order to identify the interval J which arises in the Inverse Function Rule, it is sufficient to determine the **endpoints** of J , which may be real numbers or one of the symbols ∞ and $-\infty$. For example, $(0, 1]$ has endpoints 0 and 1, and $[1, \infty)$ has endpoints 1 and ∞ . (As earlier, do not let this use of the symbol ∞ tempt you to think that ∞ is a real number.) We must also determine whether or not these endpoints *belong* to J .

Figure 40 illustrates various cases that can occur.

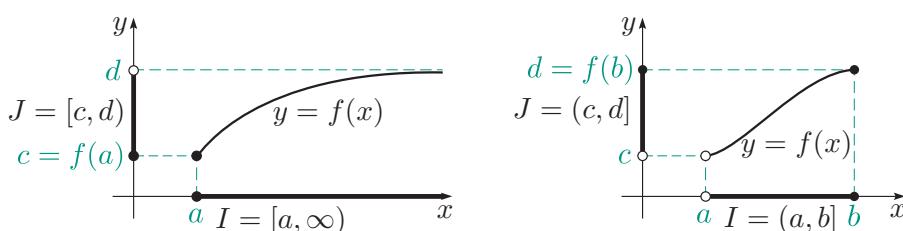


Figure 40 Two examples of the intervals I and J in the Inverse Function Rule

If a is an endpoint of I and $a \in I$, then it follows from the fact that f is increasing that $c = f(a)$ is the corresponding endpoint of J and $c \in J$.

On the other hand, if a is an endpoint of I and $a \notin I$ (this includes the possibility that a may be ∞ or $-\infty$), then it is a little harder to find the corresponding endpoint c of J . We will show in Subsection 4.4 that, in this case, if (a_n) is a monotonic sequence in I and $a_n \rightarrow a$, then $f(a_n) \rightarrow c$. This leads to the following strategy for establishing that a function has a continuous inverse function with a specified domain.

Strategy D15

To prove that $f : I \rightarrow J$, where I is an interval with endpoints a and b , has a continuous inverse function $f^{-1} : J \rightarrow I$, do the following.

1. Show that f is strictly increasing on I .
2. Show that f is continuous on I .
3. Determine the endpoint c of J corresponding to the endpoint a of I as follows:
 - if $a \in I$, then $c = f(a)$ and $c \in J$
 - if $a \notin I$, then $f(a_n) \rightarrow c$ and $c \notin J$, where (a_n) is a monotonic sequence in I such that $a_n \rightarrow a$.

Determine the endpoint d of J , corresponding to the endpoint b of I , similarly.

Note that there is a corresponding version of Strategy D15 if f is strictly decreasing. In the strictly increasing version the left endpoint of I corresponds to the left endpoint of J , whereas in the strictly decreasing version the left endpoint of I corresponds to the right endpoint of J .

Before returning to our original example, we first apply this strategy in a more straightforward case, where the domain of the function is a closed bounded interval.

Worked Exercise D53

Prove that the function

$$f(x) = x^4 + 2x + 3 \quad (x \in [0, 2])$$

has a continuous inverse function, with domain $[3, 23]$.

Solution

We use Strategy D15.

1. We showed that f is strictly increasing on $[0, \infty)$ in Exercise D57(a).

2. The function

$$f(x) = x^4 + 2x + 3 \quad (x \in [0, 2])$$

is the restriction to $[0, 2]$ of a polynomial which is continuous on \mathbb{R} . Hence f is continuous.

 We have shown that f satisfies the conditions of the Inverse Function Rule and so f has a continuous inverse function with domain $J = f([0, 2])$. We now use step 3 of the strategy to determine J by finding its endpoints. In this case, both the endpoints of the domain of f are in the domain (since it is a closed interval) and so we can find the endpoints of J by finding the images of the endpoints of $[0, 2]$. 

3. Since $f(0) = 3$ and $f(2) = 23$, we have $f([0, 2]) = [3, 23]$.

So, by the Inverse Function Rule, f has a continuous inverse function

$$f^{-1} : [3, 23] \longrightarrow [0, 2].$$

 The domain of the function is important here and we could not define an inverse function if the domain was the whole of \mathbb{R} . 

We now return to our original example.

Worked Exercise D54

Prove that the function

$$f(x) = x^5 + x - 1 \quad (x \in \mathbb{R})$$

has a continuous inverse function, with domain \mathbb{R} .

Solution

We use Strategy D15.

1. We showed that f is strictly increasing in Worked Exercise D41.
2. The function f is continuous as it is a polynomial.

 We have shown that f has a continuous inverse function with domain $f(\mathbb{R})$. We now apply step 3 of the strategy to find $f(\mathbb{R})$. In this case the interval is unbounded and so we have to find one monotonic sequence tending to ∞ and another monotonic sequence tending to $-\infty$ and then investigate the behaviour of the images of these sequences as $n \rightarrow \infty$. 

3. We first choose the increasing sequence (n) which tends to ∞ , the right endpoint of \mathbb{R} . Then

$$f(n) = n^5 + n - 1 \rightarrow \infty \text{ as } n \rightarrow \infty.$$

Thus the right endpoint of $f(\mathbb{R})$ is ∞ . We now choose the decreasing sequence $(-n)$ which tends to $-\infty$, the left endpoint of \mathbb{R} . Then

$$f(-n) = -n^5 - n - 1 \rightarrow -\infty \text{ as } n \rightarrow \infty.$$

So the left endpoint of $f(\mathbb{R})$ is $-\infty$. Thus $f(\mathbb{R}) = \mathbb{R}$.

It follows from the Inverse Function Rule that f has a continuous inverse function

$$f^{-1} : \mathbb{R} \longrightarrow \mathbb{R}.$$

The following exercise gives you the opportunity to practise using Strategy D15.

Exercise D71

Prove that the function

$$f(x) = x^2 - \frac{1}{x} \quad (x \in (0, \infty))$$

has a continuous inverse function with domain \mathbb{R} .

Hint: Use the solution to Exercise D57(b).

4.2 Inverses of standard functions

We now use the Inverse Function Rule to define continuous inverse functions for various standard functions. You are already familiar with these inverse functions, but we can now *prove* that they exist and are continuous. For each function, we give brief remarks on the three steps of Strategy D15. We also revise some of the properties of these inverse functions. Here we often use the fact that the restriction of a continuous function is continuous, mentioned at the end of Subsection 2.2.

***n*th root function**

We asserted the existence of the *n*th root function in Section 5 of Unit D1. We can now provide a proof of that result. The graph of $f(x) = x^n$ for $x \in [0, \infty)$ is shown in Figure 41.

***n*th root function**

For any positive integer $n \geq 2$, the function

$$f(x) = x^n \quad (x \in [0, \infty))$$

has a strictly increasing continuous inverse function $f^{-1}(x) = \sqrt[n]{x}$, with domain $[0, \infty)$ and image set $[0, \infty)$, called the ***n*th root function**.

We follow the steps of Strategy D15.

1. f is strictly increasing on $[0, \infty)$.
2. f is continuous on $[0, \infty)$.
3. $f(0) = 0$, and $f(k) = k^n \rightarrow \infty$ as $k \rightarrow \infty$, so

$$f([0, \infty)) = [0, \infty).$$

(We use k here, to avoid using n for two different purposes in the same expression.)

Hence f has a strictly increasing continuous inverse function

$$f^{-1} : [0, \infty) \longrightarrow [0, \infty).$$

The graph of f^{-1} is shown in Figure 42.

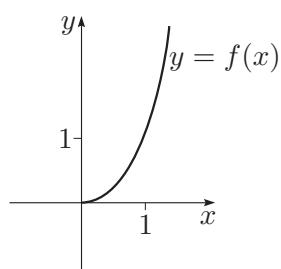


Figure 41
 $f(x) = x^n \quad (x \in [0, \infty))$

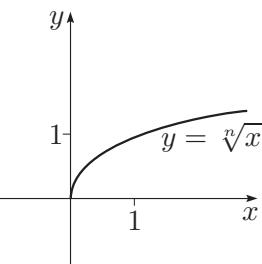


Figure 42 The *n*th root function

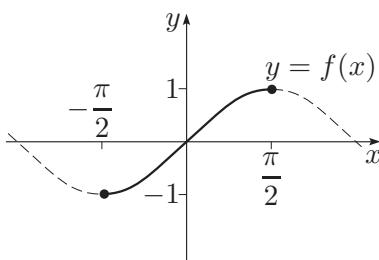


Figure 43
 $f(x) = \sin x \quad (x \in [-\pi/2, \pi/2])$

Inverse trigonometric functions

\sin^{-1}

The function

$$f(x) = \sin x \quad \left(x \in \left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right)$$

has a strictly increasing continuous inverse function, with domain $[-1, 1]$ and image set $[-\pi/2, \pi/2]$, called \sin^{-1} .

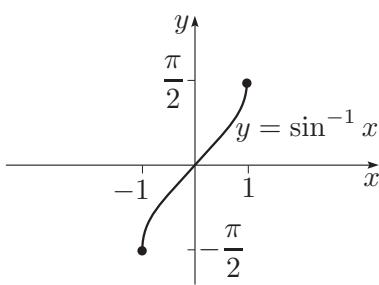


Figure 44 The inverse sine function

The graph of $f(x) = \sin x$ for $x \in [-\pi/2, \pi/2]$ is shown in Figure 43.

We follow the steps of Strategy D15.

1. The geometric definition of $f(x) = \sin x$ shows that f is strictly increasing on $[-\pi/2, \pi/2]$.
2. f is continuous on $[-\pi/2, \pi/2]$.
3. $\sin(-\pi/2) = -1$ and $\sin(\pi/2) = 1$, so

$$f\left(\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right) = [-1, 1].$$

Hence f has a strictly increasing continuous inverse function

$$f^{-1} : [-1, 1] \longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2} \right].$$

The graph of f^{-1} is shown in Figure 44.

The decreasing version of Strategy D15 can be applied similarly to prove that the cosine function has an inverse, if we restrict its domain suitably. The graph of $f(x) = \cos x$ for $x \in [0, \pi]$ is shown in Figure 45.

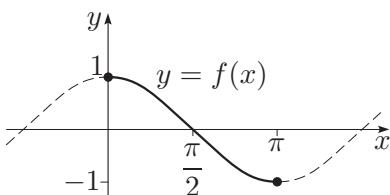


Figure 45
 $f(x) = \cos x \quad (x \in [0, \pi])$

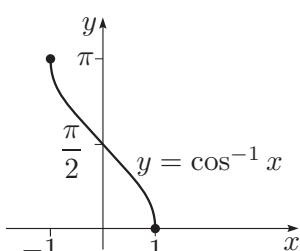


Figure 46 The inverse cosine function

\cos^{-1}

The function

$$f(x) = \cos x \quad (x \in [0, \pi])$$

has a strictly decreasing continuous inverse function, with domain $[-1, 1]$ and image set $[0, \pi]$, called \cos^{-1} .

The graph of f^{-1} is shown in Figure 46. The domain $[0, \pi]$ of f is chosen, by convention, so that f is a restriction of the cosine function which is strictly decreasing and continuous.

Similarly, to form an inverse of the tangent function, we restrict its domain to $(-\pi/2, \pi/2)$, since the tangent function is strictly increasing and continuous on this interval. The graph of the tangent function on this interval is shown in Figure 47.

tan⁻¹

The function

$$f(x) = \tan x \quad \left(x \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right) \right)$$

has a strictly increasing continuous inverse function, with domain \mathbb{R} and image set $(-\pi/2, \pi/2)$, called **tan⁻¹**.

In this case, the image set $f((-\pi/2, \pi/2))$ is \mathbb{R} because if (a_n) is a monotonic sequence in $(-\pi/2, \pi/2)$ and $a_n \rightarrow \pi/2$ as $n \rightarrow \infty$, then

$$f(a_n) = \tan a_n = \frac{\sin a_n}{\cos a_n} \rightarrow \infty \text{ as } n \rightarrow \infty.$$

The graph of f^{-1} is shown in Figure 48.

Note that some texts use arcsin, arccos and arctan instead of \sin^{-1} , \cos^{-1} , and \tan^{-1} , respectively. These names arise from the geometric definitions of the trigonometric functions.

Exercise D72

(a) Determine the values of

$$\sin^{-1}(1/\sqrt{2}), \quad \cos^{-1}(-\tfrac{1}{2}) \quad \text{and} \quad \tan^{-1}(\sqrt{3}).$$

(b) Prove that

$$\cos(2 \sin^{-1} x) = 1 - 2x^2, \quad \text{for } x \in [1, 1].$$

Hint: Let $y = \sin^{-1} x$ and use a suitable trigonometric identity from the module Handbook.

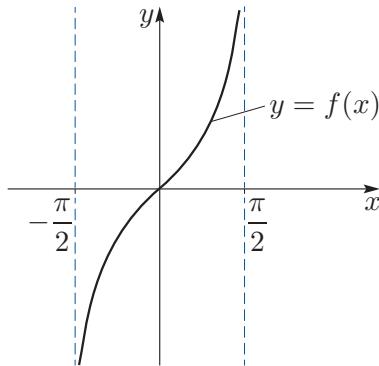


Figure 47
 $f(x) = \tan x \quad (x \in (-\pi/2, \pi/2))$

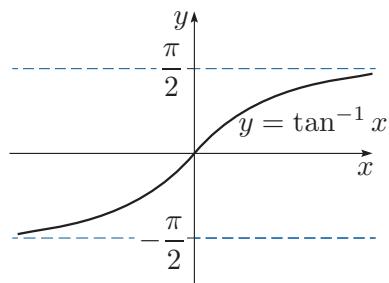


Figure 48 The inverse tangent function

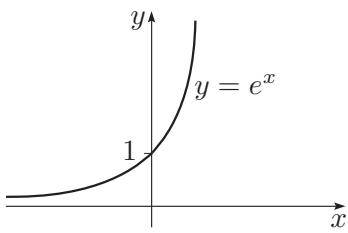


Figure 49 The exponential function

Inverse function of the exponential function

We now discuss one of the most important inverse functions: the inverse of the exponential function. The graph of the exponential function is shown in Figure 49.

log

The function

$$f(x) = e^x \quad (x \in \mathbb{R})$$

has a strictly increasing continuous inverse function f^{-1} , with domain $(0, \infty)$ and image set \mathbb{R} , called **log** or **ln**.

We follow the steps of Strategy D15.

1. f is strictly increasing on \mathbb{R} , since

$$\begin{aligned} x_1 < x_2 &\implies x_2 - x_1 > 0 \\ &\implies e^{x_2 - x_1} > 1 \quad (\text{since } e^x > 1 + x > 1, \text{ for } x > 0) \\ &\implies e^{x_2}/e^{x_1} > 1 \\ &\implies e^{x_2} > e^{x_1}. \end{aligned}$$

2. f is continuous on \mathbb{R} .

3. Since

$$f(n) = e^n \rightarrow \infty \text{ as } n \rightarrow \infty,$$

and

$$f(-n) = e^{-n} = (1/e)^n \rightarrow 0 \text{ as } n \rightarrow \infty,$$

the image set $f(\mathbb{R}) = (0, \infty)$.

Hence f has a strictly increasing continuous inverse function

$$f^{-1} : (0, \infty) \longrightarrow \mathbb{R}.$$

The graph of f^{-1} is shown in Figure 50.

Exercise D73

Prove that

$$\log(xy) = \log x + \log y, \quad \text{for } x, y \in (0, \infty).$$

Hint: Let $a = \log x$ and $b = \log y$.

Inverse hyperbolic functions

We end this subsection by considering the inverse hyperbolic functions. We first consider the sinh function which is shown in Figure 51.

sinh⁻¹

The function

$$f(x) = \sinh x = \frac{1}{2}(e^x - e^{-x}) \quad (x \in \mathbb{R})$$

has a strictly increasing continuous inverse function f^{-1} , with domain \mathbb{R} and image set \mathbb{R} , called **sinh⁻¹**.

We follow the steps of Strategy D15.

1. f is strictly increasing on \mathbb{R} , since both the functions

$$x \mapsto e^x \quad \text{and} \quad x \mapsto -e^{-x}$$

are strictly increasing on \mathbb{R} .

2. f is continuous on \mathbb{R} , by the Combination Rules.

3. Since

$$f(n) = \frac{1}{2}(e^n - e^{-n}) \rightarrow \infty \text{ as } n \rightarrow \infty,$$

and

$$f(-n) = \frac{1}{2}(e^{-n} - e^n) \rightarrow -\infty \text{ as } n \rightarrow \infty,$$

the image set $f(\mathbb{R}) = \mathbb{R}$.

Hence f has a strictly increasing continuous inverse function

$$f^{-1} : \mathbb{R} \longrightarrow \mathbb{R}.$$

The graph of f^{-1} is shown in Figure 52.

We next consider the cosh function for $x \in [0, \infty)$. The graph is shown in Figure 53.

cosh⁻¹

The function

$$f(x) = \cosh x = \frac{1}{2}(e^x + e^{-x}) \quad (x \in [0, \infty))$$

has a strictly increasing continuous inverse function f^{-1} , with domain $[1, \infty)$ and image set $[0, \infty)$, called **cosh⁻¹**.

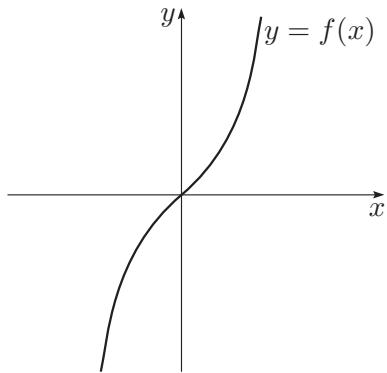


Figure 51

$f(x) = \sinh x \quad (x \in \mathbb{R})$

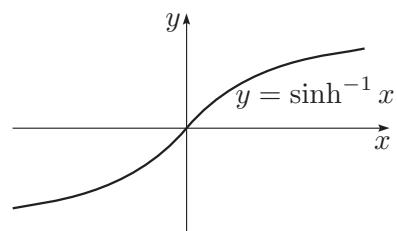


Figure 52 The inverse sinh function

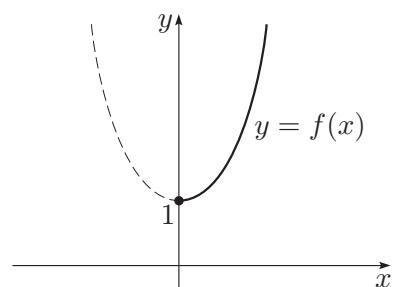


Figure 53

$f(x) = \cosh x \quad (x \in [0, \infty))$

We again follow the steps of Strategy D15.

1. f is strictly increasing on $[0, \infty)$, since $\cosh x = (1 + \sinh^2 x)^{1/2}$ and the function $x \mapsto \sinh x$ is strictly increasing on $[0, \infty)$.
2. f is continuous on $[0, \infty)$, by the Combination Rules.
3. Since $f(0) = 1$ and

$$f(n) = \frac{1}{2}(e^n + e^{-n}) \rightarrow \infty \text{ as } n \rightarrow \infty,$$

the image set $f([0, \infty)) = [1, \infty)$.

Hence f has a strictly increasing continuous inverse function

$$f^{-1} : [1, \infty) \rightarrow [0, \infty).$$

The graph of f^{-1} is shown in Figure 54.

Strategy D15 can be applied in a similar way to show that the function $f(x) = \tanh x$ is strictly increasing and continuous on \mathbb{R} , with $f(\mathbb{R}) = (-1, 1)$. The graph of this function is shown in Figure 55.

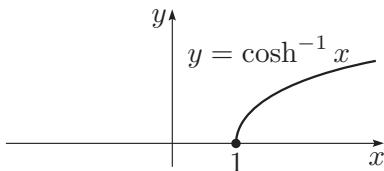


Figure 54 The inverse cosh function

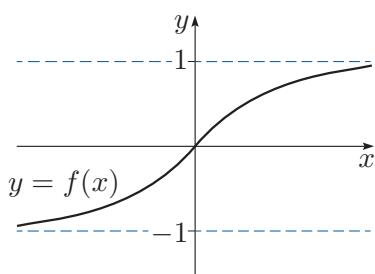


Figure 55
 $f(x) = \tanh x \quad (x \in \mathbb{R})$

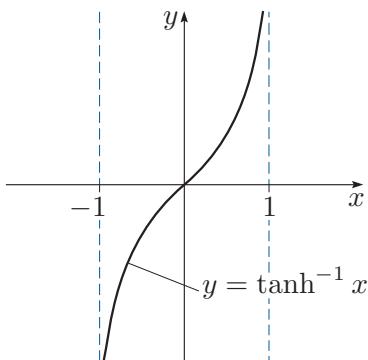


Figure 56 The inverse tanh function

tanh⁻¹

The function

$$f(x) = \tanh x = \frac{\sinh x}{\cosh x} \quad (x \in \mathbb{R})$$

has a strictly increasing continuous inverse function f^{-1} , with domain $(-1, 1)$ and image set \mathbb{R} , called **tanh⁻¹**.

The graph of the inverse tanh function is shown in Figure 56.

The inverse hyperbolic functions can all be expressed in terms of log, as we show for \sinh^{-1} in the following worked exercise.

Worked Exercise D55

Prove that

$$\sinh^{-1} x = \log \left(x + \sqrt{x^2 + 1} \right), \quad \text{for } x \in \mathbb{R}.$$

Solution

Let $y = \sinh^{-1} x$, for $x \in \mathbb{R}$, so

$$x = \sinh y = \frac{1}{2}(e^y - e^{-y}).$$

Multiplying both sides by e^y and rearranging, we obtain

$$e^{2y} - 2xe^y - 1 = (e^y)^2 - 2xe^y - 1 = 0.$$

This is a quadratic equation in e^y , with solution

$$e^y = x \pm \sqrt{x^2 + 1}.$$

Since $e^y > 0$, we must choose the + sign. We obtain

$$y = \sinh^{-1} x = \log \left(x + \sqrt{x^2 + 1} \right).$$

Exercise D74

Prove that

$$\cosh^{-1} x = \log \left(x + \sqrt{x^2 - 1} \right), \quad \text{for } x \in [1, \infty).$$

4.3 Defining exponential functions

In this subsection we consider how to properly define *exponential functions*, by which we mean functions whose rule is of the form $f(x) = a^x$ for some $a > 0$. (As you have seen, when $a = e$ we refer to the function as *the exponential function*.)

In the introduction to Book D we asked a question about the graph of $y = 2^x$, shown in Figure 57.

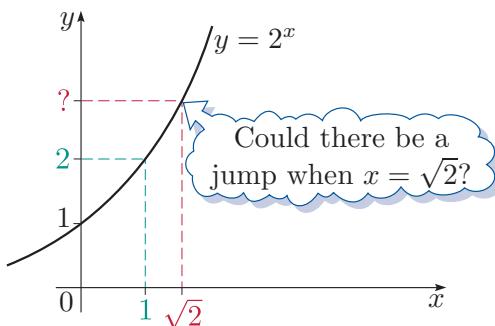


Figure 57 The graph of $y = 2^x$

In Unit D1 we defined the expression a^x for $a > 0$ when x is a rational number, but *not* when x is irrational. We now provide this missing definition, and also prove that the resulting function $x \mapsto a^x$ is continuous. In particular, it follows that the graph of $y = 2^x$ cannot have any gaps.

Recall, from Section 4 of Unit D3, that we define

$$e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n} \right)^n = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (x \geq 0)$$

and

$$e^x = (e^{-x})^{-1} \quad (x < 0).$$

As we saw in Subsection 4.2, the function $x \mapsto e^x$ is strictly increasing and continuous, and has a strictly increasing continuous inverse function

$$x \mapsto \log x \quad (x \in (0, \infty)).$$

You saw in Exercise D73 that the function \log has the property that

$$\log(ab) = \log a + \log b, \quad \text{for } a, b \in (0, \infty).$$

Thus, if $a > 0$ and $n \in \mathbb{N}$, then

$$\log(a^n) = n \log a, \quad \text{so} \quad a^n = e^{n \log a}.$$

With a little further manipulation, we can show that this equation for a^n remains true if n is replaced by any rational number x . Thus it makes sense to *define* a^x , for $a > 0$ and x irrational, by using this equation.

Definition

If $a > 0$, then

$$a^x = e^{x \log a} \quad (x \in \mathbb{R}).$$

For example,

$$2^\pi = e^{\pi \log 2}.$$

With this definition of a^x , we can verify that the function $x \mapsto a^x$ is continuous. This follows immediately from the continuity of the function $x \mapsto e^x$ and the Composition Rule. Moreover, we can also deduce the usual Index Laws for a^x from those for e^x . We state these below without proof.

Theorem D59

(a) If $a > 0$, then the function

$$x \mapsto a^x = e^{x \log a} \quad (x \in \mathbb{R})$$

is continuous.

(b) If $a, b > 0$ and $x, y \in \mathbb{R}$, then

$$a^x b^x = (ab)^x, \quad a^x a^y = a^{x+y} \quad \text{and} \quad (a^x)^y = a^{xy}.$$

In particular, it follows from Theorem D59(b) that manipulations such as

$$\left(\sqrt{2}^{\sqrt{2}}\right)^{\sqrt{2}} = \sqrt{2}^{\sqrt{2} \times \sqrt{2}} = \sqrt{2}^2 = 2 \tag{11}$$

are justified.

Equation (11) gives an unexpected proof of the result that

there exist irrational numbers a and b such that a^b is rational.

For if $\sqrt{2}^{\sqrt{2}}$ is rational, then we can take $a = b = \sqrt{2}$, but if $\sqrt{2}^{\sqrt{2}}$ is irrational, then (by equation (11)) we can take $a = \sqrt{2}^{\sqrt{2}}$ and $b = \sqrt{2}$. In fact, it can be shown using more complicated arguments that $\sqrt{2}^{\sqrt{2}}$ is irrational, but we do not prove this here.

The number $\sqrt{2}^{\sqrt{2}} = \sqrt{2^{\sqrt{2}}}$ is the square root of the number known as the Gelfond–Schneider constant $2^{\sqrt{2}}$.

In his address at the International Congress of Mathematicians in Paris in 1900, David Hilbert (1862–1943) gave his famous list of mathematical problems, the seventh of which asked for a proof that α^β is an irrational transcendental number when α is algebraic and not equal to 0 or 1, and β is irrational and algebraic. (A number is *algebraic* if it is the root of a non-zero polynomial equation with integer coefficients and *transcendental* if it is not.) He included as a particular example the number $2^{\sqrt{2}}$. In fact, Hilbert believed that proving the transcendence of $2^{\sqrt{2}}$ was such a hard problem that he said in a lecture in 1919 that he thought nobody present would live to see it proved. He turned out to be very wrong! In 1929, the Russian mathematician Aleksandr Gelfond (1906–1968) proved Hilbert's Seventh Problem for the special case where β is a quadratic irrational, which includes $2^{\sqrt{2}}$. And in 1934 Gelfond went further and obtained the general solution, as did the German mathematician Theodor Schneider (1911–1988) independently later the same year.



Aleksandr Gelfond



Theodor Schneider

Exercise D75

Use the definition of a^x to prove that each of the following functions is continuous.

- (a) $f(x) = x^\alpha$ ($x \in (0, \infty)$), where α is any fixed real number
- (b) $f(x) = x^x$ ($x \in (0, \infty)$)

4.4 Proof of the Inverse Function Rule

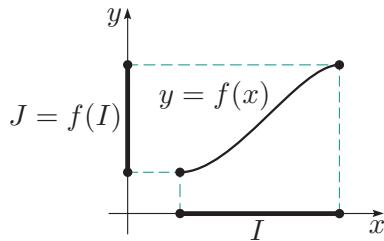


Figure 58 The graph of $f : I \rightarrow J$

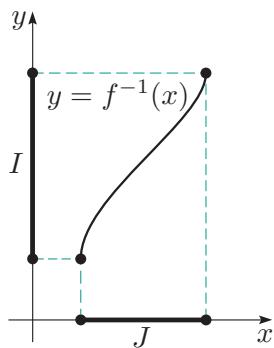


Figure 59 The graph of $f^{-1} : J \rightarrow I$

Theorem D58 Inverse Function Rule

Let $f : I \rightarrow J$, where I is an interval and J is the image set $f(I)$, be a function such that

1. f is strictly increasing on I
2. f is continuous on I .

Then J is an interval and f has an inverse function $f^{-1} : J \rightarrow I$, such that

- 1'. f^{-1} is strictly increasing on J
- 2'. f^{-1} is continuous on J .

Proof First we prove that $J = f(I)$ is an interval. Suppose that $y_1, y_2 \in f(I)$ with $y_1 < y_2$ and that y is any number in the interval (y_1, y_2) . To show that J is an interval, we must prove that $y \in f(I)$.

Now $y_1 = f(x_1)$ and $y_2 = f(x_2)$ for some $x_1, x_2 \in I$, with $x_1 < x_2$ because f is strictly increasing. Thus, since f is continuous, it follows from the Intermediate Value Theorem that there is a number $x \in (x_1, x_2)$ such that $f(x) = y$. Hence $y \in f(I)$, as required.

Next, the function f is strictly increasing and is therefore one-to-one. Thus $f^{-1} : J \rightarrow I$ exists, where $J = f(I)$.

To prove that f^{-1} is strictly increasing on J , we have to show that

$$y_1 < y_2 \implies f^{-1}(y_1) < f^{-1}(y_2), \quad \text{for } y_1, y_2 \in J.$$

💡 We use proof by contraposition. 💡

This implication holds because, for $y_1, y_2 \in J$, we have

$$\begin{aligned} f^{-1}(y_1) \geq f^{-1}(y_2) &\implies f(f^{-1}(y_1)) \geq f(f^{-1}(y_2)) \\ &\implies y_1 \geq y_2. \end{aligned}$$

Finally, we prove that f^{-1} is continuous on J . Let $y \in J$ and assume, for simplicity, that y is not an endpoint of J .

💡 Only a slight modification to the argument is needed if y is an endpoint of J . 💡

Then $y = f(x)$ for some $x \in I$, and we want to prove that

$$y_n \rightarrow y \implies f^{-1}(y_n) \rightarrow f^{-1}(y) = x.$$

Thus we assume that $y_n \rightarrow y$ and we want to deduce that

for each $\epsilon > 0$, there is an integer N such that

$$x - \epsilon < f^{-1}(y_n) < x + \epsilon, \quad \text{for all } n > N. \quad (12)$$

By taking ϵ small enough, we can assume that $x - \epsilon \in I$ and $x + \epsilon \in I$.

Now since f is strictly increasing, we know that

$$f(x - \epsilon) < f(x) < f(x + \epsilon).$$

Also, because $y_n \rightarrow y = f(x)$, there is an integer N such that

$$f(x - \epsilon) < y_n < f(x + \epsilon), \quad \text{for all } n > N.$$

Thus, since f^{-1} is a strictly increasing function, we obtain inequalities (12).

This completes the proof of the Inverse Function Rule. ■

Finally, we justify step 3 of Strategy D15 for finding the endpoints of $J = f(I)$.

Strategy D15, step 3

Let f satisfy the assumptions of the Inverse Function Rule and let a be an endpoint of I . Determine the corresponding endpoint c of J as follows:

- if $a \in I$, then $c = f(a)$ and $c \in J$
- if $a \notin I$, then $f(a_n) \rightarrow c$ and $c \notin J$, where (a_n) is a monotonic sequence in I such that $a_n \rightarrow a$.

Proof Suppose that a is the left endpoint of I ; the argument for the right endpoint is similar.

If $a \in I$, then $c = f(a) \in J$ and (since f is an increasing function)

$$f(x) \geq f(a) = c, \quad \text{for } x \in I.$$

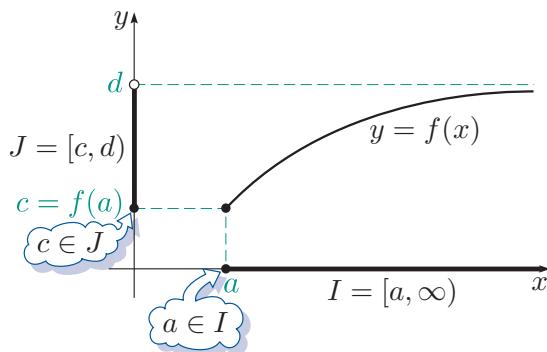


Figure 60 Step 3 of Strategy D15 when $a \in I$

Thus c is the corresponding left endpoint of J . This is illustrated in Figure 60.

If $a \notin I$, then let (a_n) be any decreasing sequence in I such that

$$a_n \rightarrow a \text{ as } n \rightarrow \infty. \quad (13)$$

Then $(f(a_n))$ is also a decreasing sequence (because f is an increasing function). Thus, by the Monotonic Sequence Theorem (Theorem D23 in Section 5 of Unit D2),

$$f(a_n) \rightarrow c \text{ as } n \rightarrow \infty, \quad (14)$$

where c is a real number or $-\infty$. This is illustrated in Figure 61.

We now prove that c is an endpoint of J and $c \notin J$.

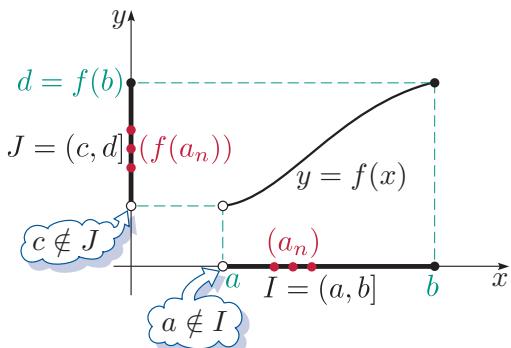


Figure 61 Step 3 of Strategy D15 when $a \notin I$

First we show that $(c, f(a_1)) \subset J$. If $c < y < f(a_1)$, then (by statement (14)) there exists n such that $f(a_n) < y < f(a_1)$, so $y = f(x)$ for some $x \in (a_n, a_1)$, by the Intermediate Value Theorem, and hence $y \in J$.

Finally, we show that $c \notin J$, using proof by contradiction. If $c \in J$, then $c = f(x)$ for some $x > a$, so (by statement (13)) there exists n such that $x > a_n$, which implies that $f(x) > f(a_n)$ and hence $f(x) > c$, a contradiction. ■

Summary

In this unit you have seen how to give a precise definition of what it means for a function to be continuous and how to use this to check whether or not a function is continuous at a point. You have also seen how to show that functions are continuous by using basic continuous functions together with rules such as the Combination Rules, the Composition Rule, the Squeeze Rule and the Glue Rule.

You have investigated some important properties of continuous functions and seen a number of applications of these properties. In particular, you have learnt how to use the Intermediate Value Theorem to locate zeros of continuous functions. Finally, you have seen how the Inverse Function Rule can be used to show that many functions have continuous inverse functions and to identify the domains of these inverse functions.

Continuity is one of the most important properties of functions and is used throughout analysis. You will meet it again when you study Book F *Analysis 2*.

Learning outcomes

After working through this unit, you should be able to:

- determine the domain and rule of the *sum*, *product*, *quotient* and *composite* of two real functions
- determine whether a given real function has an *inverse function*
- explain the meaning of the phrase ' f is *continuous* at a '
- use the Combination Rules, the Composition Rule, the Squeeze Rule and the Glue Rule for continuous functions
- recognise certain basic continuous functions
- state the Intermediate Value Theorem and use it to prove that certain equations have solutions
- determine an interval which contains all the *zeros* of a given polynomial
- state the Extreme Value Theorem and the Boundedness Theorem
- use the Inverse Function Rule to establish that a given function $f : I \longrightarrow J$ has a continuous inverse function $f^{-1} : J \longrightarrow I$
- define the inverse functions of certain standard functions
- define a^x for $a > 0$ and any $x \in \mathbb{R}$.

Solutions to exercises

Solution to Exercise D54

The domain of $f + g$ is $(-\pi/2, \pi/2)$; the rule is

$$(f + g)(x) = f(x) + g(x) = e^x + \tan x.$$

The domain of fg is $(-\pi/2, \pi/2)$; the rule is

$$(fg)(x) = f(x)g(x) = e^x \tan x.$$

The domain of f/g is $(-\pi/2, 0) \cup (0, \pi/2)$; the rule is

$$(f/g)(x) = f(x)/g(x) = e^x / \tan x = e^x \cot x.$$

Solution to Exercise D55

The domain of $f \circ g$ is

$$\begin{aligned} & \{x \in \mathbb{R} : \sin x \geq 0\} \\ &= \dots \cup [-2\pi, -\pi] \cup [0, \pi] \cup [2\pi, 3\pi] \cup \dots; \end{aligned}$$

the rule is

$$(f \circ g)(x) = \sqrt{\sin x}.$$

The domain of $g \circ f$ is

$$\{x \in [0, \infty) : \sqrt{x} \in \mathbb{R}\} = [0, \infty);$$

the rule is

$$(g \circ f)(x) = \sin \sqrt{x}.$$

Solution to Exercise D56

First we solve the equation

$$y = \frac{x+3}{x-2}$$

to obtain x in terms of y . We find that

$$y = \frac{x+3}{x-2} = 1 + \frac{5}{x-2} \iff x = 2 + \frac{5}{y-1}.$$

Thus f is one-to-one, so f has an inverse function with rule $f^{-1}(y) = 2 + 5/(y-1)$.

Now we find the image set of f , which is the domain of f^{-1} . For each $x \in (2, \infty)$, we have $5/(x-2) > 0$, so $y > 1$. Hence $f((2, \infty)) \subseteq (1, \infty)$.

Also, for each $y \in (1, \infty)$, we have $y-1 > 0$, so

$$x = 2 + \frac{5}{y-1} \in (2, \infty).$$

Thus $f((2, \infty)) \supseteq (1, \infty)$, so

$$f((2, \infty)) = (1, \infty).$$

Hence the domain of f^{-1} is $(1, \infty)$ so, adopting the usual practice of denoting the domain variable by x , we have

$$f^{-1}(x) = 2 + \frac{5}{x-1} \quad (x \in (1, \infty)),$$

which may also be written as

$$f^{-1}(x) = \frac{2x+3}{x-1} \quad (x \in (1, \infty)).$$

Solution to Exercise D57

- (a) If $0 \leq x_1 < x_2$, then $2x_1 < 2x_2$ and $x_1^4 < x_2^4$.
Hence

$$x_1^4 + 2x_1 + 3 < x_2^4 + 2x_2 + 3,$$

so f is strictly increasing, and thus one-to-one.

- (b) If $0 < x_1 < x_2$, then $1/x_1 > 1/x_2$ and $x_1^2 < x_2^2$,
so $-x_1^2 > -x_2^2$. Hence

$$\frac{1}{x_1} - x_1^2 > \frac{1}{x_2} - x_2^2,$$

so f is strictly decreasing, and thus one-to-one.

Solution to Exercise D58

- (a) $\lim_{n \rightarrow \infty} 3x_n = 6$, by the Multiple Rule for sequences.

- (b) $\lim_{n \rightarrow \infty} x_n^2 = 4$, by the Product Rule for sequences.

- (c) $\lim_{n \rightarrow \infty} 1/x_n = 1/2$, by the Quotient Rule for sequences.

Solution to Exercise D59

- (a) We guess that f is continuous at $a = 2$. The domain of f is \mathbb{R} . If (x_n) is a sequence in \mathbb{R} with $x_n \rightarrow 2$, then

$$f(x_n) = x_n^3 - 2x_n^2 \rightarrow 8 - 8 = 0 = f(2),$$

by the Combination Rules for sequences.

Hence f is continuous at $a = 2$.

(b) We guess that $f(x) = \lfloor x \rfloor$ is discontinuous at $a = 1$, since

$$f(x) = 0, \quad \text{for } 0 \leq x < 1,$$

and $f(1) = 1$. We choose

$$x_n = 1 - \frac{1}{n}, \quad n = 1, 2, \dots$$

Then $x_n \rightarrow 1$ but $f(x_n) = 0$ for $n = 1, 2, \dots$, so

$$f(x_n) \rightarrow 0 \neq 1 = f(1).$$

Hence f is discontinuous at $a = 1$.

Solution to Exercise D60

(a) Let $a \in \mathbb{R}$ and let (x_n) be a sequence in \mathbb{R} with $x_n \rightarrow a$. Then

$$f(x_n) = 1 \rightarrow 1 = f(a).$$

Hence f is continuous.

(b) Let $a \in \mathbb{R}$ and let (x_n) be a sequence in \mathbb{R} with $x_n \rightarrow a$. Then

$$f(x_n) = x_n \rightarrow a = f(a).$$

Hence f is continuous.

Solution to Exercise D61

We guess that f is continuous on $[0, \infty)$.

If we let $g(x) = x^5$ and $h(x) = |x|$, so that $h(g(x)) = |x^5|$, we can express f as a composite function $f = h \circ g$. But g is continuous (on \mathbb{R}), since it is a polynomial, and h is continuous on \mathbb{R} by Worked Exercise D45. It then follows from the Composition Rule that $f = h \circ g$ is continuous on \mathbb{R} .

Solution to Exercise D62

The function $x \mapsto x^2 + 2x + 2$ is a polynomial, and so is continuous on \mathbb{R} . The function $x \mapsto \sqrt{x}$ is continuous on $[0, \infty)$ (as we saw in Worked Exercise D46) and $x^2 + 2x + 2 = (x+1)^2 + 1 > 0$ for all $x \in \mathbb{R}$; so, by the Composition Rule for continuous functions, the function $x \mapsto \sqrt{x^2 + 2x + 2}$ is continuous on \mathbb{R} .

Next, the function $x \mapsto \frac{-3x}{x^4 + 4}$ is a rational function; hence it is continuous on its domain, which is \mathbb{R} since $x^4 + 4 \neq 0$.

Then, by the Sum Rule for continuous functions, the function

$$x \mapsto \sqrt{x^2 + 2x + 2} - \frac{3x}{x^4 + 4} \quad (= f(x))$$

is continuous on \mathbb{R} .

Solution to Exercise D63

(a) We prove that

$$f(x) = \begin{cases} x^2 \cos(1/x^2), & x \neq 0, \\ 0, & x = 0, \end{cases}$$

is continuous at 0, using the Squeeze Rule.

Since

$$-1 \leq \cos(1/x^2) \leq 1, \quad \text{for } x \neq 0,$$

and $x^2 \geq 0$, we have

$$-x^2 \leq x^2 \cos(1/x^2) \leq x^2, \quad \text{for } x \neq 0.$$

Since $f(0) = 0$, we deduce that

$$-x^2 \leq f(x) \leq x^2, \quad \text{for } x \in \mathbb{R}.$$

Thus if we take $I = \mathbb{R}$, with

$$g(x) = -x^2 \quad \text{and} \quad h(x) = x^2,$$

then

$$g(x) \leq f(x) \leq h(x), \quad \text{for } x \in I,$$

so condition 1 of the Squeeze Rule holds.

Next, $f(0) = g(0) = h(0) = 0$, so condition 2 of the Squeeze Rule is satisfied.

Finally, the functions g and h are polynomials, and so they are continuous at 0. Thus condition 3 of the Squeeze Rule is satisfied.

Hence f is continuous at 0, by the Squeeze Rule.

(b) We prove that

$$f(x) = \begin{cases} \sin(1/x), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0, \end{cases}$$

is discontinuous at 0.

According to Strategy D14, we have to find *one* sequence (x_n) such that

$$x_n \rightarrow 0 \quad \text{but} \quad f(x_n) \not\rightarrow f(0) = 0.$$

Unit D4 Continuity

We use the fact that $\sin\left(2n + \frac{1}{2}\right)\pi = 1$, for $n = 0, 1, 2, \dots$, and choose

$$x_n = \frac{1}{(2n + \frac{1}{2})\pi}, \quad n = 0, 1, 2, \dots$$

Then $x_n \rightarrow 0$ and

$$\begin{aligned} f(x_n) &= \sin(1/x_n) \\ &= \sin\left(2n + \frac{1}{2}\right)\pi = 1, \quad \text{for } n = 0, 1, 2, \dots, \end{aligned}$$

so

$$f(x_n) \not\rightarrow f(0) = 0.$$

Hence f is discontinuous at 0.

Solution to Exercise D64

We use the Glue Rule.

Let $I = \mathbb{R}$ and define the functions

$$g(x) = x^3 - 3x + 5 \quad \text{and} \quad h(x) = \frac{2x+1}{3x-2}.$$

Then f is defined on I and $1 \in I$. Also,

$$f(x) = g(x), \quad \text{for } x \in (-\infty, 1),$$

and

$$f(x) = h(x), \quad \text{for } x \in (1, \infty),$$

so condition 1 of the Glue Rule holds with $a = 1$.

Moreover, $f(1) = g(1) = h(1) = 3$, so condition 2 holds.

Finally, g and h are continuous functions, being a polynomial and a rational function respectively, and so they are both continuous at 1 since the only zero of the denominator of h is $\frac{2}{3}$. Thus condition 3 holds.

Hence f is continuous at 1, by the Glue Rule.

Solution to Exercise D65

Let $g(x) = \sin x$, $h(x) = \sqrt{x}$, and $k(x) = x^2 + 1$. Then g is continuous (by Theorem D47), h is continuous (by Worked Exercise D46), and k is continuous (being a polynomial); so

$$g(h(k(x))) = \sin(\sqrt{x^2 + 1})$$

is continuous, by the Composition Rule (applied twice). Hence

$$f(x) = k(x) + 3g(h(k(x)))$$

is continuous, by the Combination Rules.

Solution to Exercise D66

Let $g(x) = -x^2$, $h(x) = e^x$ and $k(x) = x^5 - 5x^2$.

Then g and k are both continuous as they are polynomials, and $h(x)$ is continuous by Theorem D50. Hence, by the Composition Rule and the Multiple Rule,

$$7h(g(x)) = 7e^{-x^2}$$

is continuous. Finally,

$$f(x) = k(x) + 7h(g(x))$$

is continuous by the Sum Rule.

Solution to Exercise D67

We know that c lies in $(\frac{1}{2}, 1)$, so we calculate

$$f\left(\frac{3}{4}\right) = \left(\frac{3}{4}\right)^5 + \frac{3}{4} - 1 \approx -0.0127 < 0.$$

Thus c lies in $(\frac{3}{4}, 1)$, so we calculate

$$f\left(\frac{7}{8}\right) = \left(\frac{7}{8}\right)^5 + \frac{7}{8} - 1 \approx 0.388 > 0.$$

Thus c lies in $(\frac{3}{4}, \frac{7}{8})$, so we calculate

$$f\left(\frac{13}{16}\right) = \left(\frac{13}{16}\right)^5 + \frac{13}{16} - 1 \approx 0.167 > 0.$$

Thus c lies in $(\frac{3}{4}, \frac{13}{16})$, an interval of length $\frac{1}{16}$.

Solution to Exercise D68

If $f(0) = 0$ or $f(1) = 1$, then we can take $c = 0$ or $c = 1$, respectively.

Otherwise, we have $f(0) > 0$ and $f(1) < 1$, since $0 \leq f(x) \leq 1$, for $0 \leq x \leq 1$.

We consider the function

$$g(x) = f(x) - x \quad (x \in [0, 1])$$

and show that g has a zero c in $(0, 1)$.

Now g is continuous on $[0, 1]$, by the Combination Rules. Moreover,

$$g(0) = f(0) - 0 > 0$$

and

$$g(1) = f(1) - 1 < 0.$$

Thus, by the Intermediate Value Theorem, there is a number c in $(0, 1)$ such that

$$g(c) = 0, \quad \text{so} \quad f(c) = c.$$

Solution to Exercise D69

We have

$$p(-1) = -3, \quad p(0) = 1, \quad p(1) = -1, \quad p(2) = 3,$$

so

$$p(-1) < 0 < p(0),$$

$$p(0) > 0 > p(1),$$

$$p(1) < 0 < p(2).$$

Since p is continuous, we deduce by the Intermediate Value Theorem that p has a zero in each of the intervals

$$(-1, 0), \quad (0, 1), \quad (1, 2).$$

Solution to Exercise D70

When

$$p(x) = x^5 + 3x^4 - x - 1 \quad (x \in \mathbb{R}),$$

we have

$$M = 1 + \max\{|3|, |-1|, |-1|\} = 4,$$

so all the zeros of p lie in $(-4, 4)$, by Theorem D54.

Calculating $p(n)$ for integers n in $[-4, 4]$, we obtain

n	-4	-3	-2	-1	0	1	2
$p(n)$	-253	2	17	2	-1	2	77

Thus p changes sign on each of the intervals

$$[-4, -3], \quad [-1, 0], \quad [0, 1].$$

Since p is continuous, we deduce by the Intermediate Value Theorem that p has a zero in each of the intervals

$$(-4, -3), \quad (-1, 0), \quad (0, 1).$$

Thus p has at least three zeros.

Solution to Exercise D71

We use Strategy D15.

1. We showed that $-f$ is strictly decreasing on $(0, \infty)$ in Exercise D57(b), and so f is strictly increasing.

2. The function

$$f(x) = x^2 - \frac{1}{x} = \frac{x^3 - 1}{x} \quad (x \in (0, \infty))$$

is the restriction to $(0, \infty)$ of a rational function which is continuous on $\mathbb{R} - \{0\}$. Hence f is continuous.

3. Now choose the increasing sequence (n) , which tends to ∞ , the right endpoint of $(0, \infty)$. Then

$$f(n) = n^2 - 1/n \rightarrow \infty \quad \text{as } n \rightarrow \infty,$$

by the Reciprocal Rule. Thus the right endpoint of $J = f((0, \infty))$ is ∞ .

Then choose the decreasing sequence $(1/n)$, which tends to 0, the left endpoint of $(0, \infty)$. Then

$$f(1/n) = 1/n^2 - n \rightarrow -\infty \quad \text{as } n \rightarrow \infty,$$

by the Reciprocal Rule. Thus the left endpoint of $J = f((0, \infty))$ is $-\infty$.

Hence $J = (-\infty, \infty) = \mathbb{R}$, so f has a continuous inverse function

$$f^{-1} : \mathbb{R} \longrightarrow (0, \infty),$$

by the Inverse Function Rule.

Solution to Exercise D72

(a) Since $\sin(\pi/4) = 1/\sqrt{2}$ and $\pi/4$ lies in $[-\pi/2, \pi/2]$, we have

$$\sin^{-1}(1/\sqrt{2}) = \frac{\pi}{4}.$$

Since $\cos(2\pi/3) = -\frac{1}{2}$ and $2\pi/3$ lies in $[0, \pi]$,

$$\cos^{-1}(-\frac{1}{2}) = \frac{2\pi}{3}.$$

Since $\tan(\pi/3) = \sqrt{3}$ and $\pi/3$ lies in $(-\pi/2, \pi/2)$,

$$\tan^{-1}(\sqrt{3}) = \frac{\pi}{3}.$$

(b) Following the hint, we put $y = \sin^{-1} x$. Then

$$\cos(2 \sin^{-1} x) = \cos(2y) = 1 - 2 \sin^2 y = 1 - 2x^2,$$

since $x = \sin y$.

Solution to Exercise D73

Following the hint, we put $a = \log x$ and $b = \log y$. Then $x = e^a$ and $y = e^b$, so

$$\begin{aligned}\log(xy) &= \log(e^a e^b) \\ &= \log(e^{a+b}) = a + b = \log x + \log y.\end{aligned}$$

Now the functions

$$\begin{aligned}x &\mapsto \log x \quad (x \in (0, \infty)), \\ x &\mapsto x \quad (x \in \mathbb{R}), \\ x &\mapsto e^x \quad (x \in \mathbb{R}),\end{aligned}$$

are all continuous, so f is continuous by the Product Rule and the Composition Rule.

Solution to Exercise D74

Let $y = \cosh^{-1} x$, where $x \geq 1$. Then

$$x = \cosh y = \frac{1}{2}(e^y + e^{-y}).$$

Hence, by multiplying both sides by e^y ,

$$e^{2y} - 2xe^y + 1 = (e^y)^2 - 2xe^y + 1 = 0.$$

This is a quadratic equation in e^y , with solutions

$$e^y = x \pm \sqrt{x^2 - 1}.$$

Both choices of \pm give a positive expression on the right, but we also have $e^y \geq 1$, since
 $y = \cosh^{-1} x \geq 0$.

Since

$$(x + \sqrt{x^2 - 1})(x - \sqrt{x^2 - 1}) = x^2 - (x^2 - 1) = 1$$

and $x + \sqrt{x^2 - 1} \geq 1$, we have $x - \sqrt{x^2 - 1} \leq 1$.

Thus we choose the $+$ sign, to give

$$y = \cosh^{-1} x = \log \left(x + \sqrt{x^2 - 1} \right).$$

(The value $y = \log \left(x - \sqrt{x^2 - 1} \right)$ gives the negative solution of the equation $\cosh y = x$.)

Solution to Exercise D75

(a) For $x > 0$, we have

$$f(x) = x^\alpha = e^{\alpha \log x}.$$

Now the functions

$$\begin{aligned}x &\mapsto \log x \quad (x \in (0, \infty)), \\ x &\mapsto e^x \quad (x \in \mathbb{R}),\end{aligned}$$

are both continuous, so f is continuous by the Multiple Rule and the Composition Rule.

(b) For $x > 0$, we have

$$f(x) = x^x = e^{x \log x}.$$

Acknowledgements

Grateful acknowledgement is made to the following sources.

Cover image: © Mark Owen

Unit D3

Jean le Rond d'Alembert (Subsection 2.1): The Granger Collection / Universal Images Group

Unit D4

Stanislaw Ulam (Subsection 3.3): Unless otherwise indicated, this information has been authored by an employee or employees of the Los Alamos National Security, LLC (LANS), operator of the Los Alamos National Laboratory under Contract No. DE-AC52-06NA25396 with the U.S. Department of Energy. The U.S. Government has rights to use, reproduce, and distribute this information. The public may copy and use this information without charge, provided that this Notice and any statement of authorship are reproduced on all copies. Neither the Government nor LANS makes any warranty, express or implied, or assumes any liability or responsibility for the use of this information.

Theodor Schneider (Subsection 4.3): Taken from wikipedia.org and used under <https://creativecommons.org/licenses/by-sa/2.0/deed.en>

Every effort has been made to contact copyright holders. If any have been inadvertently overlooked the publishers will be pleased to make the necessary arrangements at the first opportunity.

Index

Absolute Convergence Test 161
 absolute value 20
 absolutely convergent 160
 Alternating Test 165
 strategy for using 166
 Antipodal Points Theorem 233
 Apéry, Roger 149
 Archimedean algorithm 115
 Archimedean Property 11
 Archimedes 112
 arithmetic in \mathbb{R} 13, 49

basic continuous functions 221
 basic null sequences 86
 basic series 157
 Bernoulli's Inequality 37
 Bernoulli, Jacob 36, 116
 binomial coefficient 31
 Binomial Theorem 32, 86
 bisection method 224
 Bolzano, Bernard 222
 Borsuk, Karol 234
 bounded above
 for a sequence 110
 for a set 40
 bounded below
 for a sequence 110
 for a set 42
 bounded closed interval 221
 bounded function 230
 bounded sequence 100
 bounded set 42
 Boundedness Theorem 231

Cauchy, Augustin-Louis 75, 222
 codomain convention 187
 combination of functions 188
 Combination Rules
 for continuous functions 208
 for convergent series 138
 for inequalities 24
 for sequences 81, 91, 105
 Comparison Test 147
 strategy for using 148
 composite function 191

Composition Rule
 for continuous functions 209
 constant sequence 67, 89
 continuous
 at a point 201
 on a set 201
 convergent sequence 89, 90, 101
 convergent series 131
 general strategy 168
 \cos^{-1} 240
 \cosh^{-1} 243

d'Alembert, Jean le Rond 156
 de Sluse, René-François 36
 de Stainville, Janot 174
 decimal 5, 6
 decreasing function 195
 decreasing sequence 67
 Dedekind cut 14
 Density Property 11
 digit 5
 discontinuous 201
 divergent sequence 100
 strategy for proving 108
 divergent series 131
 domain convention 187
 dominant term 93
 dominated 84

e 115
 endpoint 236
 Euler, Leonhard 116, 147, 174
 even subsequence 106
 eventual property of a sequence 72
 exponential function e^x 169, 219
 inverse 242
 exponential function a^x 245
 Exponential Inequalities 219
 extreme value 230
 Extreme Value Theorem 231

field 12
 finite decimal 5
 First Subsequence Rule 108
 Fourier, Joseph 174

Gelfond, Aleksandr 247

- geometric series
 - finite 130
 - infinite 136
- Glue Rule
 - for continuous functions 214
- greatest lower bound of a set 45
 - strategy for checking 46
- Greatest Lower Bound Property of \mathbb{R} 47
- harmonic series 145
- Harriot, Thomas 116
- Hilbert, David 247
- Hui, Yang 33
- hybrid function 212–215
- hyperbolic function
 - inverse 243–244
- image set 188, 197
- increasing 195
- increasing sequence 67
- Index Laws 53, 246
- inequalities
 - proofs 23
 - rearranging 15
 - rules for rearranging 21
 - solving 17
- infimum of a set 45
- infinite decimal 6
- infinite series 131
- integer 4
- Intermediate Value Theorem 222
- inverse function 192
 - exponential 242
 - hyperbolic 243–244
 - trigonometric 240–241
- Inverse Function Rule 235
- irrational number 9
- irrationality of e 173
- Khayyam, Omar 33
- least upper bound of a set 43
 - strategy for checking 44
- Least Upper Bound Property of \mathbb{R} 46
- Leibniz Test *see* Alternating Test
- Limit Comparison Test 150
 - strategy for using 152
- Limit Inequality Rule 98
- limit of a sequence 89
- ln 242
- local property of continuity 216
- log 242
- logarithm 242
- lower bound of a set 42
 - greatest 45, 47
- mathematical induction 35
- maximum element of a set 40
- maximum/minimum value of a function 230
- Mengoli, Pietro 147
- minimum element of a set 42
- modulus 20
 - properties of 21
- Monotone Convergence Theorem 110
- monotonic function 195
- monotonic sequence 67
 - strategies for proving 70
- Monotonic Sequence Theorem 111
- multiple of a function 189
- Multiple Rule
 - for continuous functions 208
 - for convergent series 138
 - for divergent series 140
 - for sequences 81, 91, 105
- natural number 4
- Newton, Isaac 33
- non-convergent sequence *see* divergent sequence
- Non-null Test 141
 - strategy for using 142
- non-recurring decimal 9
- non-terminating decimal 6
- n th partial sum of a series 131
- n th term of a series 131
- n th root 51, 239
- n th term of a sequence 64
- null sequence 76
 - basic 86
 - strategy for proving 79, 84
- odd subsequence 106
- one-to-one function 188, 194–196
- onto 188
- order properties of \mathbb{R} 4, 11
- Oresme, Nicole 145
- Pascal's triangle 33

- Pascal, Blaise 33
 π 112
 Power Rule 80
 powers 52
 product of functions 189
 Product Rule
 for continuous functions 208
 for inequalities 24
 for sequences 81, 91, 94, 105
 proving inequalities 23
 quotient of functions 189
 Quotient Rule
 for continuous functions 208
 for sequences 91, 95
 Ratio Test 154
 rational number 4
 real function 187
 real line 10
 real number 9
 Reciprocal Rule for sequences 103
 recurring decimal 6
 restriction of a function 217
 root 50
 rules for rearranging inequalities 21
 Schneider, Theodor 247
 Second Subsequence Rule 108
 sequence 64
 bounded 100
 constant 67, 89
 convergent 89, 90, 101
 decreasing 67
 divergent 100, 108
 increasing 67
 limit of 89
 monotonic 67, 70
 nth term of 64
 null 76, 79, 84, 86
 strictly decreasing 67
 strictly increasing 67
 tends to infinity 102
 unbounded 100
 sequence diagram 66
 sigma notation 134
 Sine Inequality 217
 \sin^{-1} 240
 \sinh^{-1} 243
 solution set 17
 solving inequalities 17
 Squeeze Rule
 for continuous functions 210
 for sequences 84, 96, 105
 strict inequality 16
 strictly increasing/decreasing function 195
 strictly increasing/decreasing sequence 67
 strictly monotonic 195
 subsequence 106
 even 106
 odd 106
 Subsequence Rules 108
 sum of functions 189
 sum of series 131
 Sum Rule
 for continuous functions 208
 for convergent series 138
 for inequalities 24
 for sequences 81, 91, 94, 105
 supremum of a set 43
 \tan^{-1} 241
 \tanh^{-1} 244
 telescoping series 137
 terminating decimal 5
 Transitive Property 11
 Transitive Rule
 for inequalities 23
 Triangle Inequality 24
 Triangle Inequality (infinite form) 163
 Trichotomy Property 11
 trigonometric function 217–219
 inverse 240–241
 Ulam, Stanislaw 234
 unbounded sequence 100
 unbounded set 42
 upper bound of a set 40
 least 43
 vanish (of a function) 226
 Wallis, John 33
 weak inequality 16
 Zeno's paradoxes 129
 zero of a function 226

